Email : bilgumus@gmail.com
Visit our home page on www.dergipark.org.tr/mejs

Tamraz H. TAMRAZOV
Field: Biological Sciences
(Assoc.Prof.Dr., Department of Plant Physiology and Biotechnology, Research Institute of Crop Husbandry, Ministry of Agriculture of the Republic of Azerbaijan)
tamraz.tamrazov@mail.ru

Yusuf ZEREN

Field: Mathematics, Topology

(Assoc.Prof.Dr., Yıldız Technical University, Faculty of Science and Letters, Dep. of Mathematics, TURKEY)

yzeren@yildiz.edu.tr

Z. Gökay KAYNAK

Field: Nuclear Physics

(Retired Prof.Dr., Uludag University, Faculty of Science, Dep. of Physics, Turkey) kaynak@uludag.edu.tr

Middle East Journal of Science

https://dergipark.org.tr/mejs

e-ISSN:2618-6136

MEJS
Volume 10
Issue 1
June 2024

CONTENTS

Research Articles

Review Article

Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Research Article

ON SOME INEQUALITIES FOR EXPONENTIALLY WEIGHTED FRACTIONAL HARDY OPERATORS WITH Δ-INTEGRAL CALCULUS

Lütfi AKIN*[1]

[1]Department of Business Administration, Mardin Artuklu University, Mardin/Turkiye
[*] Corresponding author; lutfiakin@artuklu.edu.tr

Abstract: *Dynamic equations, inequalities, and operators are the indispensable cornerstones of harmonic analysis and time-scale calculus. Undoubtedly, one of the most important of these operators and inequalities is the Hardy operator and inequality. Because especially when we say variable exponent Lebesgue space, the first thing that comes to our mind is the Hardy operator. We know that the topics in question have many applications in different scientific fields. In this paper, some inequalities will be proved for variable exponentially weighted Hardy operators with Δ-integral calculus.*

Keywords: *Hardy inequality, Variable exponent, Weight function, Time scale.*

Received: March 11, 2024 *Accepted: May 06, 2024*

1. Introduction

The variable exponent Lebesgue space was first revealed by [1]. However, this space is based on the paper of [2] together with applications to modeling electrorheological fluids [3]. Many operators and inequalities have been studied in the variable exponent Lebesgue space. One of the most important of these is the Hardy operator, fractional Hardy operator, and inequality. As an example of a few studies, Different approaches have been introduced to the Hardy operator, fractional Hardy operator, and inequalities [4, 5]. In addition, a new dimension was added to the Hardy-type inequalities in the weighted and variable exponent Lebesgue space [6]. Mathematicians and scientists working in different scientific fields have expanded the workspace of operators in harmonic analysis. For example, they have proved the conditions for the boundedness and compactness, etc. of operators and fractional inequalities in variable exponent Lebesgue spaces and they used it in physics, mechanics, electrorheological fluids, optics, economics, etc. [7, 8, 9, 10, 11, 12, 13]. Besides, it has been known that the fractional Hardy operator in a variable exponent Lebesgue space does not satisfy arbitrary non-negative measurable functions; but provides for non-negative monotonic functions. Moreover, the sharp constant of the fractional Hardy-type operator was obtained for non-negative functions [13]. After a while, the monotony was replaced by a weaker condition [14]. However, a relationship has also been established between harmonic analysis in general, the variable exponential Lebesgue space in particular, and other spaces of different types. Time scales, Morrey spaces, and Sobolev spaces can be examples of spaces.

Recently, the calculation of time scales has also attracted authors' attention. Generally, by integrating time scales with the subjects within the field of harmonic analysis, they have revealed the magnificent relationships between them. For example, the dynamic integral inequalities and fractional integral operators have been studied on time scales with variable exponent Lebesgue spaces by many authors [15-28].

We aim of this paper is to obtain some inequalities for exponentially weighted fractional Hardy-type operator and the weighted dual of the classical fractional Hardy operator acting exponentially weighted with Δ-integral in time scale calculus.

2. Auxiallary Statements and Preliminaries

The emergence of the theory of time scales was introduced to the literature in 1988 by Stefan Hilger [29]. Although it later experienced a period of stagnation, its popularity increased especially after the 2000s. Recently, many authors have studied certain dynamic inequalities on time scales, operators, and concepts that fall within the field of harmonic analysis. We see that it still maintains its popularity today. Let $[x, y]$ be a facultative closed interval on $\mathbb{T}$ (time scale). We refer to the references [29-32] for more details. $[x, y]_{\mathbb{T}}$ is denoted by $[x, y] \cap \mathbb{T}$.

Definition 2.1. [31] The functions $\sigma, \rho: \mathbb{T} \to \mathbb{T}$ are defined by $\sigma(t) = inf\{s \in \mathbb{T}: s > t\}$, $\rho(t) = sup\{s \in \mathbb{T}: s < t\}$ for $t \in \mathbb{T}$. $\sigma(t)$ is defined as forward jump operator. The function $\rho(t)$ is defined as backward jump operator. If $\sigma(t) > t$, then t is defined as right-scattered. If $\sigma(t) = t$, then t is called as right-dense. If $\rho(t) < t$, then t is defined as left-scattered. If $\rho(t) = t$, then t is called as left-dense.

Definition 2.2. [31] Let functions $\mu, \vartheta: \mathbb{T} \to \mathbb{R}^+$ such that $\mu(t) = \sigma(t) - t,\ \vartheta(t) = t - \rho(t)$. The functions $\mu(t)$ and $\vartheta(t)$ are called as graininess functions. Let $\mathbb{T}$ be a left-scattered maximum m, then $\mathbb{T}^k = \mathbb{T} - \{m\}$.

$\mathbb{T}^k$ is defined as follows

$$\mathbb{T}^k = \begin{cases} \mathbb{T} \setminus (\rho\, sup\, \mathbb{T}, sup\, \mathbb{T}], & if \quad sup\, \mathbb{T} < \infty \\ \mathbb{T}, & if \quad sup\, \mathbb{T} = \infty \end{cases}$$

and

$$\mathbb{T}_k = \begin{cases} \mathbb{T} \setminus [inf\, \mathbb{T}, \sigma(inf\, \mathbb{T})], & |inf\, \mathbb{T}| < \infty \\ \mathbb{T}, & inf\, \mathbb{T} = -\infty. \end{cases}$$

Let $g: \mathbb{T} \to \mathbb{R}$ be a mapping and let t be defined as right-dense. We can write the following.

i) Let g be Δ −differentiable at $t \in \mathbb{T}^k (t \neq min\mathbb{T})$, then g is continuous at point t.

ii) Let g be left continuous at t and let t be defined as right-scattered, then g is Δ −differentiable at point t,

$$g^{\Delta}(t) = \frac{g^{\sigma}(t) - g(t)}{\mu(t)}$$

iii) Let g be Δ −differentiable at t and $\lim\limits_{s \to t} \frac{g(t)-g(s)}{t-s}$, then

$$g^{\Delta}(t) = \lim_{s \to t} \frac{g(t) - g(s)}{t - s}.$$

iv) Let g be Δ −differentiable at t, then $g^{\sigma}(t) = g(t) + \mu(t) g^{\Delta}(t)$.

Remark 2.3. [31] If $\mathbb{T} = \mathbb{R}$, then $g^{\Delta}(t) = g'(t)$, and if $\mathbb{T} = \mathbb{Z}$, then $h^{\Delta}(t) = \Delta h(t)$.

Definition 2.4. [30] If $G: \mathbb{T} \to \mathbb{R}$ is defined as a Δ −antiderivative of $g: \mathbb{T} \to \mathbb{R}$, then $G^{\Delta} = g(t)$ holds for all $t, s \in \mathbb{T}$ and Δ-integral of g is called as the by

$$\int_{s}^{t} g(\tau)\Delta\tau = G(t) - G(s).$$

The following statements were proved by [14]

Theorem 2.5. [14] Let f be a measurable function on $(0,\infty)$ and satisfies for some $C > 0$. Let $x > 0$, $0 < q < 1$, and $\beta < 1 - \frac{1}{q}$. The inequality

$$f(x) \leq \frac{C}{x}\left(\int_{0}^{x} f^{q}(t)t^{q-1}dt\right)^{1/q}, \tag{1}$$

then

$$\left\|x^{\beta}(Hf)(x)\right\|_{L_q(0,\infty)} \leq K\left\|y^{\beta}f(y)\right\|_{L_q(0,\infty)} \tag{2}$$

where

$$K = C^{1-q}q^{1-\frac{1}{q}}\left(1-\beta-\frac{1}{q}\right)^{-\frac{1}{q}} \tag{3}$$

Moreover, the constant K is sharp.

Let ω denote a weight non-negative function. The space of $L_{p,\omega}(0,\infty)$ for $0 < p < \infty$ is defined as follows:

$$\|f\|_{L_{p,\omega}(0,\infty)} = \left(\int_{0}^{\infty} |f(x)|^{p}\omega(x)dx\right)^{1/p}.$$

The fractional Hardy operator is defined as follows:

$$(H_{\omega}f)(x) = \frac{1}{K(x)}\frac{1}{x^{1-\propto}}\int_{0}^{x} f(y)\omega(y)dy, \quad x > 0,$$

where $0 \leq \propto < 1$, $0 < K(x) = \int_0^x \omega(y)dy < \infty$ for all $y > 0$.

Lemma 2.6. [14] Let $0 < q < 1$, $k_1 > 0$, $B > 0$, and let ω be a non-negative weight function, then the following inequality can be written.

$$\omega(s) \leq k_1\omega(t) \quad \text{for} \quad 0 < t < s < \infty \tag{4}$$

Let f be a measurable function for almost all $0 < s < \infty$, then

$$f(s) \leq B\left(\int_{0}^{s} \omega(t)t^{q-1}dt\right)^{-1/q}\left(\int_{0}^{s} f^{q}(t)\omega(t)t^{q-1}dt\right)^{1/q} \tag{5}$$

for all $x > 0$

$$(H_{\omega}f)(x) \leq \frac{k_2}{x\omega(x)^{1/q}}\left(\int_{0}^{x} f^{q}(t)\omega(t)t^{q-1}dt\right)^{1/q} \tag{6}$$

where $k_2 = q^{1/q}B^{1-q}c_1^{\frac{z}{q}-1}$.

Remark 2.7. [14] If $\omega = 1$, then inequality (5) turns into inequality (1) with $C = Bp^{1/p}$ and $k_1 = 1$, consequently $k_2 = p^{1/p}B^{1-p}$.

Remark 2.8. [14] If f is non-increasing mapping, then (5) holds for $B = 1$.

Theorem 2.9. [14] Let $k_1 > 0, 0 < q < 1, B > 0$, ω be a positive weight function, and $\beta < 1 - \frac{1}{q}$. Let f be a Lebesgue non-negative measurable function, then

$$\left\| x^{\beta}(H_{\omega}f)(x) \right\|_{L_{q,\omega}(0,\infty)} \leq N \left\| y^{\beta} f(y) \right\|_{L_{q,\omega}(0,\infty)} \tag{7}$$

where

$$N = B^{1-q} c_1^{\frac{2}{q}-1} \left(1 - \beta - \frac{1}{q}\right)^{-\frac{1}{q}} \tag{8}$$

Let H_{ω}^{*} be the dual of the operator H_{ω} in $L_2(0,\infty)$. Then for any $f, g \in L_2(0,\infty)$

$$\int_0^{\infty} \left(\frac{1}{D(x)} \int_0^{x} f(s)\omega(s)ds \right) g(x)dx = \int_0^{\infty} \left(\int_y^{\infty} \frac{g(x)}{D(x)} dx \right) f(s)\omega(s)ds$$

$$= \int_0^{\infty} \omega(s)(H^{*}g)(x)f(s)ds = \int_0^{\infty} \omega(y) \left(\int_y^{\infty} \frac{g(x)}{D(x)} dx \right) f(s)ds.$$

Hence the equality $(H_{\omega}f, g)_{L_2(0,\infty)} = (f, H_{\omega}^{*}g)_{L_2(0,\infty)}$ is satisfied for the operator H_{ω}^{*} defined by

$$(H_{\omega}^{*}f)(x) = \omega(x) \int_x^{\infty} \frac{g(y)}{D(s)} ds, \quad x > 0.$$

Lemma 2.10. Let ω be a non-negative weight function for $x > 0$, and let $\int_0^x \omega(t)dt < \infty$ be satisfied. Let f be a non-negative measurable function for $0 < x < \infty$. Let $0 < q < 1, B > 0$,

$$\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt < \infty,$$

and

$$f(x) \leq \frac{B}{x} \left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt \right)^{1/q} \omega(x)^{\frac{1}{1-q}} \left(\int_0^x \omega(t)dt \right)^{\frac{1}{1-q}}, \tag{9}$$

then for $n > 0$

$$(H_{\omega}^{*}f)(x) \leq k_3 \omega(x) \left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt \right)^{1/q} \tag{10}$$

where $k_3 = qB^{1-q}$.

Proof. By (9) it follows that

$$x^{1-q}f(x)^{1-q} \leq B^{1-q} \left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt \right)^{\frac{1}{q}-1} \omega(\mathrm{x}) \int_0^x \omega(t)dt.$$

Hence

$$\frac{f(x)}{K(x)} \leq B^{1-q} \omega(\mathrm{x}) f(x)^q x^{q-1} \left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt \right)^{\frac{1}{q}-1}$$

$$= qB^{1-q}(-1)\left[\left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt\right)^{1/q}\right].$$

Integrating over (n,∞), then we get

$$\int_n^{\infty} \frac{f(x)}{D(x)}dx$$

$$\leq qB^{1-q}\lim_{x\to\infty}\left(\left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt\right)^{q} - \left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt\right)^{1/q}\right)$$

$$\leq qB^{1-q}\left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt\right)^{1/q}$$

Hence

$$(H_\omega^* f)(x) = \omega(x)\int_x^{\infty}\frac{f(x)}{D(x)}dx \leq qB^{1-q}\omega(x)\left(\int_x^{\infty} f^q(t)\omega(t)t^{q-1}dt\right)^{1/q}.$$

If $\omega(x) = 1$ in (9) and (10), then we have the following corollary.

Corollary 2.11. [14] Let f be a non-negative Lebesgue measurable function for $0 < x < \infty$, and $\int_x^{\infty} f^q(t)t^{q-1}dt < \infty$. Let $B > 0$ and $0 < q < 1$, then the inequality

$$f(x) \leq \frac{B}{x^{q'}}\left(\int_x^{\infty} f^q(t)t^{q-1}dt\right)^{1/q} \tag{11}$$

is satisfied, then for $x > 0$

$$(H^* f)(x) \leq k_3\left(\int_x^{\infty} f^q(t)t^{q-1}dt\right)^{1/q} \tag{12}$$

where $k_3 = qB^{1-q}$ and q' is the conjugate exponent of q.

Remark 2.12. Inequalites (11), (12) respectively are analogues of inequality (1) and inequality (2) in [8], for the dual of the classical Hardy operator.

Theorem 2.13. [14] Let f be a non-negative Lebesgue measurable function. Let $-\frac{1}{q} < \beta < 1 - \frac{1}{q}$, $0 < q < 1$, $B > 0, x > 0$, then

$$\left\|\gamma^\beta (H^* f)(\gamma)\right\|_{L_q(0,\infty)} \leq k_4\left\|t^{\beta+1}f(t)\right\|_{L_q(0,\infty)} \tag{13}$$

where $k_4 = qB^{1-q}\ (\beta q + 1)^{-\frac{1}{q}}$.

Proof.

$$L_1 = \left\|\gamma^\beta (H^* f)(\gamma)\right\|_{L_q(0,\infty)} = \left[\int_0^{\infty}\gamma^{\beta q}(H^* f)^q(\gamma)d\gamma\right]^{1/q} = \left[\int_0^{\infty}\gamma^{\beta q}\left(\int_\gamma^{\infty}\frac{f(t)}{t}dt\right)^q d\gamma\right]^{1/q}.$$

Then it follows that

$$L_1 \leq \left[\int_0^{\infty}\gamma^{\beta q}k_3^q\left(\int_\gamma^{\infty} f^q(t)t^{q-1}dt\right)^q d\gamma\right]^{1/q} = k_3\left[\int_\gamma^{\infty} f^q(t)t^{q-1}\left(\int_0^{t}\gamma^{\beta q}d\gamma\right)dt\right]^{1/q}$$

$$= qB^{1-q} \ (\beta q + 1)^{-\frac{1}{q}} \left\| t^{\beta+1} f(t) \right\|_{L_q(0,\infty)}.$$

Let φ be a measurable positive function in R^m. Suppose that p is a measurable positive function on φ. Assume that $0 < p^- \leq p(x) \leq p^+ < \infty$, $p^- = ess \inf_{x \in \varphi} p(x)$, $p^+ = ess \sup_{x \in \varphi} p(x)$ and ω is a weight function on φ.

Definition 2.14. Let $L_{p(s),\omega}(\varphi)$ we define as all measurable functions on φ such that

$$I_{p,\omega}(f) = \int_{\varphi} (|f(s)\omega(s)|)^{p(s)} ds < \infty \tag{14}$$

Note that the expression

$$\|f\|_{L_{p(.),\omega}(\varphi)} = \inf\left\{\lambda > 0;\ \int_{\varphi} \left(\frac{|f(s)|\omega(s)}{\lambda}\right)^{p(s)} ds \leq 1\right\} \tag{15}$$

denotes on $L_{p(s),\omega}(\varphi)$.

Corollary 2.15. [10] Let $n(s) = \frac{u(s)v(s)}{v(s)-u(s)}$, and let $0 < u^- \leq u(s) \leq v(s) \leq v^+ < \infty$. Assume that ω_1, ω_2 are weight functions in φ satisfying the condition:

$$\left\|\frac{\omega_1}{\omega_2}\right\|_{L_{n(.)}(\varphi)} < \infty.$$

Then the inequality

$$\|f\|_{L_{u(.),\omega_1}(\varphi)} \leq \left(A_1 + B_1 + \|\chi\varphi_2\|_{L_\infty(\varphi)}\right)^{\frac{1}{u^-}} \left\|\frac{\omega_1}{\omega_2}\right\|_{L_{n(.)}(\varphi)} \|f\|_{L_{v(.),\omega_2}(\varphi)} \tag{16}$$

holds for every $f \in L_{v(s),\omega_2}(\varphi)$, where

$$\varphi_1 = \{s \in \varphi : u(s) < v(s)\}, \quad \varphi_2 = \{y \in \varphi : u(s) = v(s)\},$$

$$B_1 = \sup_{s \in \varphi_1} \frac{u(s)}{v(s)}, \quad A_1 = \sup_{s \in \varphi_1} \frac{v(s) - u(s)}{v(s)}.$$

Lemma 2.16. [11] Let $t \in \varphi_2 \subset R^m$. If $0 < p^- \leq p(x) \leq q(t) \leq q^+ < \infty$, for all $x \in \varphi_1 \subset R^m$, and if $p \in M(\varphi_1)$, then the inequality

$$\left\| \|f\|_{L_{p(.)}(\varphi_1)} \right\|_{L_{q(.)}(\varphi_2)} \leq M_{p,q} \left\| \|f\|_{L_{q(.)}(\varphi_2)} \right\|_{L_{p(.)}(\varphi_1)} \tag{17}$$

holds, where

$$M_{p,q} = \left(\|\chi\Lambda_1\|_\infty + \|\chi\Lambda_2\|_\infty + \frac{p^+}{q^-} + \frac{p^-}{q^+}\right)(\|\chi\Lambda_1\|_\infty + \|\chi\Lambda_2\|_\infty) \tag{18}$$

$$q^- = ess \inf_{\varphi_2} q(x), \ q^- = ess \sup_{\varphi_2} q(x)$$

$$\Lambda_1 = \{(x,t) \in \varphi_1 \times \varphi_2; p(x) = q(x)\}, \qquad \Lambda_2 = \varphi_1 \times \varphi_2 \backslash \Lambda_1.$$

If $M(\varphi_1)$ is define as continuous functions on φ_1 and if $f: \varphi_1 \times \varphi_2 \to R$ is define as measurable function, then

$$\left\| \|f\|_{L_{q(.)}(\varphi_2)} \right\|_{L_{p(.)}(\varphi_1)} < \infty.$$

Theorem 2.17. [10] Let f be a non-negative and non-increasing function, and let $0 < p^- \leq p(x) \leq q(t) \leq q^+ < 1$, for $x \in (0,\infty)$, and equality $n(x) = \frac{p^- p(x)}{p(x)-p^-}$ be satisfied. Suppose that φ_1 and φ_2 are weight non-negative functions. The inequality

$$\|Hf\|_{L_{q(x),\omega_2}(0,\infty)} \leq p^{-\frac{1}{p^-}} M_{p,q} b_p \left\| \frac{y^{\frac{1}{p^+}} \left\| \frac{\omega_2}{x} \right\|_{L_{q(.)}(y,\infty)}}{\omega_1} \right\|_{L_{n(.)}(0,\infty)} \|f\|_{L_{p(.),\omega_1}(0,\infty)} \qquad (19)$$

holds for $f \in L_{p(x),\omega_1}(0,\infty)$, where

$$M_{p,q} = \left(\|\chi\Lambda_1\|_{L_\infty(0,\infty)} + \|\chi\Lambda_2\|_{L_\infty(0,\infty)} + p^-\left(\frac{1}{q^-} - \frac{1}{q^+}\right)\right)\left(\|\chi\mathrm{T}_1\|_{L_\infty(0,\infty)} + \|\chi\mathrm{T}_2\|_{L_\infty(0,\infty)}\right),$$

$\mathrm{T}_1 = \{x \in (0,\infty): p(x) = p^-\}$, $\mathrm{T}_2 = (0,\infty)\backslash\mathrm{T}_1$, and $b_p = \left(1 - \frac{p^+ - p^-}{p^+} + \|\chi\mathrm{T}_1\|_{L_\infty(0,\infty)}\right)^{\frac{1}{p^-}}$.

3. Results and Discussion

The variable exponentially fractional Hardy-type operator with Δ −integral in time scale calculus is defined by as follows

$$(H_{\propto\omega^p} f)(x) = \frac{1}{K(x)} \frac{1}{x^{1-\propto}} \int_0^x f(s)\omega^{p(s)}(s)\Delta s, \quad x > 0,$$

where $0 \leq \propto < 1$, $0 < p(s) < 1$, $0 < K(x) = \int_0^x \omega^{p(s)}(s)\Delta s < \infty$ for all $s > 0$.

Theorem 3.1. Let f is Δ −integrable non-negative Lebesgue measurable function satisfying inequality (5) with p replaced by p^-. ω is a positive weight function. Let $0 \leq \propto < 1$, $0 < p^- \leq p(x) \leq q(x) \leq q^+ < 1$, $\beta < 1 - \frac{1}{p^-}$, $n(x) = \frac{p^- p(x)}{p(x)-p^-}$. Assume that ω_1 and ω_2 are weight positive functions. The inequality

$$\|H_{\propto\omega^{p(t)}} f\|_{L_{q(x),\omega_2}(0,\infty)}$$

$$\leq k_2 M_{p,q} b_p \left\| \frac{\omega^{\frac{p(t)}{p^-}} t^{\frac{1}{p^-}} \left\| \frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)} \right\|_{L_{q(.)}(t,\infty)}}{\omega_1^{p(y)}} \right\|_{L_{n(.)}(0,\infty)} . \|f\|_{L_{p(.),\omega_1^{p(t)}}(0,\infty)} \qquad (20)$$

is valid for $f \in L_{p(x),\omega_1}(0,\infty)$, where $k_2 = p^{-\frac{1}{p^-}} k_1^{\frac{2}{p^-}-1} B^{1-p^-}$.

Proof. Applying Lemma 2.6, we obtain

$$\|H_{\propto\omega^{p(y)}} f\|_{L_{q(x),\omega_2^{p(y)}}(0,\infty)} = \left\|\omega_2^{p(y)} H_{\propto\omega^{p(y)}} f\right\|_{L_{q(x)}(0,\infty)}$$

$$\leq \left\| \frac{k_2\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\left(\frac{1}{x^{1-\propto}}\int_0^x f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}$$

$$= k_2\left\| \frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\left(\frac{1}{x^{1-\propto}}\int_0^x f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}.$$

Let $J_1 = \left\| \frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\left(\frac{1}{x^{1-\propto}}\int_0^x f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}$

$$J_1 = \left\|\left(\frac{1}{x^{1-\propto}}\int_0^\infty \left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(0.x)}(t)\left[\frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\right]^{p^-} t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}$$

$$= \left\|\left(\frac{1}{x^{1-\propto}}\int_0^\infty \left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(0.x)}(t)\left[\frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\right]^{p^-} t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}$$

$$= \left\|\left\|\frac{1}{x^{1-\propto}}\left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(0.x)}(t)\left[\frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\right]^{p^-} t^{p^--1}\right\|_{L_1(0,\infty)}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}^{\frac{1}{p^-}}.$$

Next applying Lemma 2.16, we obtain

$$J_1 \leq M_{p,q}\left(\frac{1}{x^{1-\propto}}\int_0^\infty \left\|\left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(0.x)}(t)\left[\frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\right]^{p^-} t^{p^--1}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}\Delta t\right)^{1/p^-}$$

$$= M_{p,q}\left(\frac{1}{x^{1-\propto}}\int_0^\infty f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\left\|\chi_{(0.x)}(t)\left[\frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\right]^{p^-}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}\Delta t\right)^{1/p^-}$$

$$= M_{p,q}\left(\frac{1}{x^{1-\propto}}\int_0^\infty f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\left\|\left[\frac{\omega_2^{p(t)}(x)}{x\omega^{\frac{p(t)}{p^-}}(x)}\right]^{p^-}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}\Delta t\right)^{1/p^-}$$

$$= M_{p,q} \left\| \frac{1}{x^{1-\propto}} f(t) \omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \frac{\omega_2^{p(t)}(x)}{x \omega^{\frac{p(t)}{p^-}}(x)} \right\|_{L_q(t,\infty)} \right\|_{L_{p^-}(0,\infty)}.$$

Let $J_2 = \left\| f(t) \omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \frac{\omega_2^{p(t)}(x)}{x \omega^{\frac{p(t)}{p^-}}(x)} \right\|_{L_q(t,\infty)} \right\|_{L_{p^-}(0,\infty)}$,

then applying Corollary 2.15, we obtain

$$J_2 \leq b_p \left\| \frac{\omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \frac{\omega_2^{p(t)}(x)}{x \omega^{\frac{p(t)}{p^-}}(x)} \right\|_{L_q(t,\infty)}}{\omega_1^{p(t)}} \right\|_{L_{n(.)}(0,\infty)} \frac{1}{x^{1-\propto}} \|f\|_{L_{p(.)\omega_1^{p(t)}}(0,\infty)},$$

Hence

$$\|H_{\omega^p} f\|_{L_{q(x),\omega_2}(0,\infty)}$$

$$\leq k_2 M_{p,q} b_p \left\| \frac{\omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \frac{\omega_2^{p(t)}(x)}{x \omega^{\frac{p(t)}{p^-}}(x)} \right\|_{L_q(t,\infty)}}{\omega_1^{p(t)}} \right\|_{L_{n(.)}(0,\infty)} \frac{1}{x^{1-\propto}} \|f\|_{L_{p(.)\omega_1^{p(t)}}(0,\infty)}.$$

Theorem 3.2. Let f be $\Delta -$integrable a non-negative Lebesgue measurable function satisfying inequality (9) with p replaced by p^-. ω is a positive weight function. Let $0 \leq \propto < 1$, $0 < p^- \leq p(x) \leq q(x) \leq q^+ < 1$, $\beta < 1 - \frac{1}{p^-}$, $n(x) = \frac{p^- p(x)}{p(x) - p^-}$, $x \in (0,\infty)$. Assume that ω_1 and ω_2 are positive weight functions. The inequality

$$\left\| H^*_{\propto \omega^{p(t)}} f \right\|_{L_{q(x),\omega_2^p}(0,\infty)}$$

$$\leq k_3 M_{p,q} b_p \left\| \frac{\omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \omega_2^{p(t)}(x) \omega^{p(t)}(x) \right\|_{L_q(0,t)}}{\omega_1^{p(t)}} \right\|_{L_{n(.)}(0,\infty)} \|f\|_{L_{p(.)\omega_1^{p}(t)(0,\infty)}} \tag{21}$$

is valid for $f \in L_{p(x),\omega_1}(0,\infty)$, where $k_3 = p^- A^{1-p^-}$.

Proof. Applying Lemma 2.6, we obtain

$$\left\| H^*_{\propto \omega^{p(t)}} f \right\|_{L_{q(x),\omega_2^p(t)}(0,\infty)} = \left\| \omega_2^{p(t)} H^*_{\omega^{p(t)}} f \right\|_{L_{q(x)}(0,\infty)}$$

$$\leq k_3 \left\| \omega_2^{p(t)}(x)\omega^{p(t)}(x)\left(\frac{1}{x^{1-\propto}}\int\limits_x^\infty f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}$$

Let $V_1 = \left\| \omega_2^{p(t)}(x)\omega^{p(t)}(x)\left(\frac{1}{x^{1-\propto}}\int_x^\infty f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}$,

$$V_1 = \left\| \left(\frac{1}{x^{1-\propto}}\int\limits_0^\infty \left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(x,\infty)}(t)\left[\omega_2^{p(t)}(x)\omega^{p(t)}(x)\right]^{p^-} t^{p^--1}\Delta t\right)^{1/p^-}\right\|_{L_{q(x)}(0,\infty)}$$

$$= \left\| \frac{1}{x^{1-\propto}}\int\limits_0^\infty \left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(x,\infty)}(t)\left[\omega_2^{p(t)}(x)\omega^{p(t)}(x)\right]^{p^-} t^{p^--1}\Delta t\right\|^{1/p^-}_{L_{\frac{q(.)}{p^-}}(0,\infty)}$$

$$= \left\| \left\| \frac{1}{x^{1-\propto}}\left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(x,\infty)}(t)\left[\omega_2^{p(t)}(x)\omega^{p(t)}(x)\right]^{p^-} t^{p^--1}\right\|_{L_1(0,\infty)}\right\|^{1/p^-}_{L_{\frac{q(.)}{p^-}}(0,\infty)}$$

Applying Lemma 2.16, we obtain

$$V_1 \leq M_{p,q}\left(\frac{1}{x^{1-\propto}}\int\limits_0^\infty \left\| \left[f^{p^-}(t)\omega^{p(t)}(t)\right]\chi_{(x,\infty)}(t)\left[\omega_2^{p(t)}(x)\omega^{p(t)}(x)\right]^{p^-} t^{p^--1}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}\Delta t\right)^{1/p^-}$$

$$= M_{p,q}\left(\frac{1}{x^{1-\propto}}\int\limits_0^\infty f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\left\| \chi_{(x,\infty)}(t)\left[\omega_2^{p(t)}(x)\omega^{p(t)}(x)\right]^{p^-}\right\|_{L_{\frac{q(.)}{p^-}}(0,\infty)}\Delta t\right)^{1/p^-}$$

$$= M_{p,q}\left(\frac{1}{x^{1-\propto}}\int\limits_0^\infty f^{p^-}(t)\omega^{p(t)}(t)t^{p^--1}\left\| \omega_2^{p(t)}(x)\omega^{p(t)}(x)\right\|^{p^-}_{L_q(0,t)}\Delta t\right)^{1/p^-}$$

$$= M_{p,q}\left\| \frac{1}{x^{1-\propto}}f(t)\omega^{\frac{p(t)}{p^-}}(t)t^{\frac{p^--1}{p^-}}\left\| \omega_2^{p(t)}(x)\omega^{p(t)}(x)\right\|_{L_q(0,t)}\right\|_{L_{p^-}(0,\infty)}.$$

By applying Corollary 2.15, we obtain

$$\left\| \frac{1}{x^{1-\propto}}f(t)\omega^{\frac{p(t)}{p^-}}(t)t^{\frac{p^--1}{p^-}}\left\| \omega_2^{p(t)}(x)\omega^{p(t)}(x)\right\|_{L_q(0,t)}\right\|_{L_{p^-}(0,\infty)}$$

$$\leq b_p \left\| \frac{\omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \omega_2^{p(t)}(x) \omega^{p(t)}(x) \right\|_{L_q(0,t)}}{\omega_1^{p(t)}} \right\|_{L_{n(.)}(0,\infty)} \frac{1}{x^{1-\propto}} \|f\|_{L_{p(.)\omega_1^{p(t)}}(0,\infty)}.$$

Hence

$$\left\| H^*_{\omega^{p(t)}} f \right\|_{L_{q(x),\omega_2^{p}}(0,\infty)}$$

$$\leq k_3 M_{p,q} b_p \left\| \frac{\omega^{\frac{p(t)}{p^-}}(t) t^{\frac{p^- -1}{p^-}} \left\| \omega_2^{p(t)}(x) \omega^{p(t)}(x) \right\|_{L_q(0,t)}}{\omega_1^{p(t)}} \right\|_{L_{n(.)}(0,\infty)} \frac{1}{x^{1-\propto}} \|f\|_{L_{p(.)\omega_1^{p(t)}}(0,\infty)}.$$

Remark 3.3. In Theorem 3.1 and Theorem 3.2, if we get $\mathbb{T} = \mathbb{R}$, $\propto = 1$ and $p(t) = 0$, then we obtain continuous weighted inequalities as mentioned in [15].

4. Conclusion

In general, operators and variable exponent types of inequalities have become one of the important cornerstones of harmonic analysis. Their boundedness, compactness, etc. have caused them to become the focus of attention of mathematicians. Numerous studies have been conducted in this field, especially since the beginning of this century. Likewise, the issue of time scales has become popular, especially in the last years, although not so much in the past. Holistic studies covering time scales and harmonic analysis have become more popular recently. Inspired by these studies, we took a new present to variable exponentially fractional Hardy operator by reconciling time scales and harmonic analysis.

Ethical Statement

The author declares that this document does not require ethics committee approval or any special permission.

Conflict of Interest

The author declares no conflict of interest.

References

[1] Orlicz, W. "Uber konjugierte Exponentenfolgen," *Stud. Math*., 3, 200-212, 1931.

[2] Kovacik, O., Rakosnik,J. "On spaces $L^{p(x)}$ and $W^{k,p(x)}$," *Czechoslovak Math. J*., 41, no. 4, 592-618, 1991.

[3] Akın, L. "A Characterization of Boundedness of Fractional Maximal Operator with Variable Kernel on Herz-Morrey Spaces." *Anal. Theory Appl*., Vol. 36, No. 1, pp. 60-68, 2020.

[4] Akin, L. "A Characterization of Approximation of Hardy Operators in VLS", *Celal Bayar University Journal of Science,* Volume 14, Issue 3, pp:333-336, 2018.

[5] Akın, L., Zeren, Y. “On innovations of the multivariable fractional Hardy-type inequalities on time scales”. *Sigma J Eng Nat Sci* ;41(2):415−422, 2023.

[6] Bandaliev, R.A. “On Hardy-type inequalities in weighted variable exponent spaces $L^p(x)$ for $0 < p < 1$,” *Eurasian Math.J*., 4, no. 4, 5-16, 2013.

[7] Mamedov, F.I., Zeren, Y., Akin, L. “Compactification of weighted Hardy operator in variable exponent Lebesgue spaces,” *Asian Journal of Mathematics and Computer Research*, 17(1): 38-47, 2017.

[8] Azzouz, N., Halim, B., Senouci, A. “An inequality for the weighted Hardy operator for $0 < p < 1$”, *Eurasian Math. J*., 4, no. 3, 60-65, 2013.

[9] Ruzicka, M. Electrorheological Fluids: Modeling and Mathematical Theory, Lecture Notes in Mathematics, Springer, Berlin. 1748, 2000.

[10] Bandaliev, R.A. “On an inequality in Lebesgue space with mixed norm and with variable summability exponent,” *Mat. Zametki*, 84, no. 3, 323-333 (in Russian). English translation: Math. Notes, 84,no. 3, 303-313, 2008.

[11] Samko, S.G. “Differentiation and integration of variable order and the spaces $L^p(x)$,” Proc. Inter. Conf. "Operator theory for complex and hypercomplex analysis", Mexico, 1994, Contemp. Math., 212, 203-219, 1998.

[12] Burenkov, V.I. “Function spaces. Main integral inequalities related to L^p-space,” Peoples' Friendship University of Russia, Moscow, 96pp, 1989.

[13] Burenkov, V.I. “On the exact constant in the Hardy inequality with $0 < p < 1$ for monotone functions,” Trudy Matem. Inst. Steklov 194, 58-62 (in Russian), 1992.; English translation in proc. Steklov Inst. Math., 194, no. 4, 59-63, 1993.

[14] Senouci, A., Tararykova, T. “Hardy-type inequality for $0 < p < 1$,” *Evraziiskii Matematicheskii Zhurnal*, pp.112-116, 2007.

[15] Bendaoud, S.A., Senouci, A. “Inequalities for weighted Hardy operators in weighted variable exponent Lebesgue space with $0 < p(x) < 1$,” *Eurasian Math. J.,* Volume 9, Number 1, 30–39, 2018.

[16] Akın, L. “On some results of weighted Hölder type inequality on time scales,” *Middle East Journal of Science*. 6(1), 15-22 2020.

[17] Akın, L. “On innovations of n-dimensional integral-type inequality on time scales.” *Adv. Differ. Equ.* 148 (2021), 2021.

[18] Akın, L. “On the Fractional Maximal Delta Integral Type Inequalities on Time Scales,” *Fractal Fract.* 4(2), 1-10, 2020.

[19] Agarwal, R.P., Bohner, M., Saker, S.H. “Dynamic Littlewood-type inequalities”. *Proc. Am. Math. Soc.* 143(2), 667–677, 2015.

[20] Oguntuase, J.A., Persson, L.E. “Time scales Hardy-type inequalities via superquadracity.” *Ann. Funct. Anal*. 5(2), 61–73, 2014.

[21] Rehak, P. “Hardy inequality on time scales and its application to half-linear dynamic equations.” *J. Inequal. Appl.* 5, 495–507, 2005.

[22] Saker, S.H. “Hardy–Leindler type inequalities on time scales.” *Appl. Math. Inf. Sci*. 8(6), 2975–2981, 2014.

[23] Saker, S.H., O’Regan, D. “Extensions of dynamic inequalities of Hardy’s type on time scales.” *Math. Slovaca* 65(5), 993–1012, 2015.

[24] Saker, S.H., O’Regan, D. “Hardy and Littlewood inequalities on time scales.” *Bull. Malays. Math. Sci. Soc*. 39(2), 527–543, 2016.

[25] Saker, S.H., O’Regan, D., Agarwal, R.P. “Some dynamic inequalities of Hardy’s type on time scales.” *Math. Inequal. Appl.*17, 1183–1199, 2014.

[26] Saker, S.H., O’Regan, D., Agarwal, R.P. “Generalized Hardy, Copson, Leindler and Bennett inequalities on time scales.” *Math. Nachr*. 287(5–6), 686–698, 2014.

[27] Saker, S.H., O’Regan, D., Agarwal, R.P. “Dynamic inequalities of Hardy and Copson types on time scales.” *Analysis 3*4,391–402, 2014.

[28] Saker, S.H., O’Regan, D., Agarwal, R.P. “Littlewood and Bennett inequalities on time scales.” *Mediterr. J. Math*. 12, 605–619, 2015.

[29] Hilger, S. Ein Maßkettenkalkül mit Anwendung auf Zentrmsmannigfaltingkeiten, Ph.D. Thesis, Univarsi. Würzburg, 1988.

[30] Saker, S.H., Rezk, H.M., Krni´c, M. “More accurate dynamic Hardy-type inequalities obtained via superquadraticity.” *Rev. R. Acad. Cienc. Exactas Fís. Nat., Ser. A Mat*. 113(3), 2691–2713, 2019.

[31] Saker, S.H., Saied, A.I., Krni´c, M. “Some new weighted dynamic inequalities for monotone functions involving kernels.” *Mediterr. J. Math*. 17(2), 1–18, 2020.

[32] Saker, S.H., Saied, A.I., Krni´c, M. “Some new dynamic Hardy-type inequalities with kernels involving monotone functions.” *Rev. R. Acad. Cienc. Exactas Fís. Nat., Ser. A Mat.* 114, 1–16, 2020.

Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Research Article

ON ρ −STATISTICAL CONVERGENCE OF SEQUENCES OF FUZZY NUMBERS

Damla BARLAK*

Dicle University, Faculty of Science, Department of Statistics, 21280, Diyarbakır, Turkiye
* Corresponding author; dyagdiran@hotmail.com

Abstract: *In this study, we present the concepts of ρ −statistically convergence for sequences of fuzzy numbers as well as strong $\left(w_\rho(F)\right)$ summability and ρ −Cauchy statistically convergence for sequences of fuzzy numbers. We also provide several results concerning these concepts.*

Keywords: *Cesàro summability, Statistical convergence, Strongly ρ −Cesàro summability.*

Received: June 4, 2024 *Accepted: June 28, 2024*

1. Introduction

Fast [1] gave short description of statistical convergence 1951. Schoenberg [2] investigated statistical convergence as a summability method and outlined several fundamental properties associated with it. This concept has been applied by many researchers under different names to measurement theory, locally convex spaces, summability theory, Banach spaces, trigonometric series in Fourier analysis and theory of fuzzy set ([3],[4],[5],[6]). The concept of statistical convergence depend on the density subsets of the set $\mathbb{N}$. The natural density of a subset K of $\mathbb{N}$ is defined by $\delta(K) = \lim_{n\to\infty} \frac{1}{n}|\{k \le n: k \in K\}|$, if the limit exists, where the vertical bars indicate number of the elements in $\{k \le n: k \in K\}$.

If x is a sequence such that satisfies feature P for all k apart from a set of naturally density zero, then we say that x_k satifies P for "almost all k" and we shortened this by "$a.a.k$."

Fuzzy set theory, which is a very valuable logic with accuracy, was first introduced by Zadeh [7] in 1965. The applications of this theory span various fields, including fuzzy topological spaces, fuzzy measurements, fuzzy mathematical programming, and fuzzy logic. The concept of fuzzy number sequence is first encountered in Matloka's paper [8].

Matloka [8] defined the concept of bounded and convergent sequences of fuzzy numbers and studied their some properties. Since then, many studies on sequences of fuzzy numbers have been made and studies on this subject are still ongoing ([9], [10], [11], [12], [13], [14]).

A fuzzy number is fuzzy set $Z: \mathbb{R} \to [0,1]$ with the following properties:

$i)$ Z is normal, that is, there exists an $z_0 \in \mathbb{R}$ such that $Z(z_0) = 1$;

$ii)$ Z is fuzzy convex, that is, for $z, t \in \mathbb{R}$ and $0 \le \lambda \le 1, Z(\lambda z + (1-\lambda)t) \ge \min[Z(z), Z(t)]$;

$iii)$ Z is upper semicontinuous;

$iv)$ $suppZ = cl\{z \in \mathbb{R}: Z(z) > 0\}$, or denoted by $[Z]^0$, is compact.

The definition α −level set $[Z]^\alpha$ of a fuzzy number is determined by

$$[Z]^{\alpha} = \begin{cases} \{z \in \mathbb{R}: Z(z) \geq \alpha\}, & \text{if } \alpha \in (0,1] \\ suppZ, & \text{if } \alpha = 0. \end{cases}$$

It is evident that Z is a fuzzy number is necessary and sufficient for $[Z]^{\alpha}$ is a closed interval for each $\alpha \in [0,1]$ and $[Z]^1 \neq \emptyset$. The set of all fuzzy number sequences will be denoted as $L(\mathbb{R})$. The distance between two fuzzy numbers Z and T, we use the metric

$$d(Z,T) = \sup_{0 \leq \alpha \leq 1} d_H([Z]^{\alpha}, [T]^{\alpha})$$

Let $Z^{\alpha} = \left[\underline{Z}^{\alpha}, \overline{Z}^{\alpha}\right]$ and $T^{\alpha} = \left[\underline{T}^{\alpha}, \overline{T}^{\alpha}\right]$. Then, the Hausdorff metric is characterized by

$$d_H([Z]^{\alpha}, [T]^{\alpha}) = \max\left\{\left|\underline{Z}^{\alpha} - \underline{T}^{\alpha}\right|, \left|\overline{Z}^{\alpha} - \overline{T}^{\alpha}\right|\right\}.$$

It is known that d is a metric on $L(\mathbb{R})$, and $(L(\mathbb{R}), d)$ is a complete metric space.

Nuray and Savaş [15] defined the concept of statistical convergence for sequences of fuzzy numbers. A sequence $Z = (Z_k)$ of fuzzy numbers is a function $Z: \mathbb{N} \to L(\mathbb{R})$. Let $Z = (Z_k)$ be a sequence of fuzzy numbers. Then the sequence $Z = (Z_k)$ fuzzy numbers, is called statistically convergent to the fuzzy number Z_0 if for each $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{1}{n} |\{k \leq n: d(Z_k, Z_0) \geq \varepsilon\}| = 0.$$

The set of all fuzzy number sequences demonstrating statistically convergent will be denoted as $S(F)$.

Çakallı [16] defined the concept of $\rho -$statistically convergence. Subsequently, many authors have done a great deal of work on $\rho -$ statistical convergence([17],[18],[19],[20],[21],[22]). The aim of this paper is to extend the investigation conducted by Çakallı [16].

2. Main Results

In this section, we present the concepts of $\rho -$statistically convergence for sequences of fuzzy numbers, strong $\left(w_{\rho}(F)\right)$ summability for sequences of fuzzy numbers and $\rho -$Cauchy statistically convergence for sequences of fuzzy numbers. We also provide several results pertaining to these concepts.

Definition 2.1. Let (Z_k) be a fuzzy number sequence, the sequence $Z = (Z_k)$ is is called $\rho -$statistically convergent to the fuzzy number Z_0 if

$$\lim_{n\to\infty} \frac{1}{\rho_n} |\{k \leq n: d(Z_k, Z_0) \geq \varepsilon\}| = 0$$

for each $\varepsilon > 0$, where $\rho = (\rho_n)$ is a non-decreasing sequence for each $n \in \mathbb{Z}^+$ tending to ∞ such that $\limsup_n \frac{\rho_n}{n} < \infty$, $\Delta\rho_n = O(1)$ and $\Delta Z_n = Z_{n+1} - Z_n$ for each $n \in \mathbb{Z}^+$.

In this case, either $S_{\rho}(F) - \lim Z_k = Z_0$ or $Z_k \to Z_0 \left(S_{\rho}(F)\right)$ is used as a notation. The set of all fuzzy number sequences demonstrating ρ-statistical convergence will be denoted as $S_{\rho}(F)$. If for each

$n \in \mathbb{N}$ $\rho = (\rho_n) = n$, the concept of being $\rho-$statistically convergent is equivalent to being statistically convergent.

Definition 2.2. Let (Z_k) be a fuzzy number sequence, the sequence $Z = (Z_k)$ is is called strong $\rho-$convergent (or $\left(w_\rho(F)\right)-$convergent) to Z_0 if

$$\lim_{n\to\infty} \frac{1}{\rho_n} \sum_{k=1}^{n} d(Z_k, Z_0) = 0.$$

In this case, either $\left(w_\rho(F)\right) - \lim Z_k = Z_0$ or $\lim_{k\to\infty} Z_k \to Z_0 \left(w_\rho(F)\right)$ is used as a notation. The set of all fuzzy number sequences demonstrating strong $\rho-$convergent will be denoted as $\left(w_\rho(F)\right)$.

Theorem 2.1 Let (Z_k) and (T_k) be two fuzzy numbers sequences, $\rho = (\rho_n)$ is a non-decreasing sequence for each $n \in \mathbb{Z}^+$ tending to ∞ such that $\limsup_n \frac{\rho_n}{n} < \infty$, $\Delta\rho_n = O(1)$ and $\Delta Z_n = Z_{n+1} - Z_n$ for each $n \in \mathbb{Z}^+$. Then

(i) $Z_k \to Z_0 \left(S_\rho(F)\right)$ and $c \in \mathbb{C}$ implies $(cZ_k) \to cZ_0 \left(S_\rho(F)\right)$,

(ii) $Z_k \to Z_0 \left(S_\rho(F)\right)$ and $T_k \to T_0 \left(S_\rho(F)\right)$ implies $(Z_k + T_k) \to (Z_0 + T_0) \left(S_\rho(F)\right)$.

Proof. (i) For $c = 0$, the proof is clear. Let $c \neq 0$, the inequality leads to the proof

$$\frac{1}{\rho_n} |\{k \leq n : d(cZ_k, cZ_0) \geq \varepsilon\}| \leq \frac{1}{\rho_n} \left|\left\{k \leq n : d(Z_k, Z_0) \geq \frac{\varepsilon}{c}\right\}\right|.$$

(ii) Let $Z_k \to Z_0 \left(S_\rho(F)\right)$ and $T_k \to T_0 \left(S_\rho(F)\right)$, we can write

$$\frac{1}{\rho_n} |\{k \leq n : d(Z_k + T_k, Z_0 + T_0) \geq \varepsilon\}|$$

$$\leq \frac{1}{\rho_n} \left|\left\{k \leq n : d(Z_k, Z_0) \geq \frac{\varepsilon}{2}\right\}\right| + \frac{1}{\rho_n} \left|\left\{k \leq n : d(T_k, T_0) \geq \frac{\varepsilon}{2}\right\}\right|$$

for each $\varepsilon > 0$ and thus if $Z_k \to Z_0 \left(S_\rho(F)\right)$ and $T_k \to T_0 \left(S_\rho(F)\right)$ then $(Z_k + T_k) \to (Z_0 + T_0) \left(S_\rho(F)\right)$.

Definition 2.3 Let (Z_k) be a fuzzy number sequence, the sequence $Z = (Z_k)$ is called $S_\rho(F)-$Cauchy sequence if there exists a subsequence $\left(Z_{k'(n)}\right)$ of Z such that $k'(n) \leq n$ for every n, $\lim_{n\to\infty} Z_{k'(n)} = Z_0$ and for each $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{1}{\rho_n} \left|\left\{k \leq n : d\left(Z_k, Z_{k'(n)}\right) \geq \varepsilon\right\}\right| = 0,$$

where $\rho = (\rho_n)$ is non-decreasing sequence for each $n \in \mathbb{Z}^+$ tending to ∞ such that $\limsup_n \frac{\rho_n}{n} < \infty$, $\Delta\rho_n = O(1)$ and $\Delta Z_n = Z_{n+1} - Z_n$ for each $n \in \mathbb{Z}^+$.

Theorem 2.2. The subsequent statements are mutually equivalent:

(i) (Z_k) is a ρ −statistical convergence,

(ii) (Z_k) is a ρ −Cauchy statistical convergence,

(iii) (Z_k) is a sequence of fuzzy numbers for which there is a ρ −statistically convergent sequence of fuzzy numbers T such that $Z_k = T_k\ a.a.k$.

Theorem 2.3 Let (Z_k) be a fuzzy number sequence, the sequence $Z = (Z_k)$ is $S_\rho(F)$ −convergent a necessary and sufficient condition is that (Z_k) is an $S_\rho(F)$ −Cauchy sequence.

Proof. Let's consider Z_k is an S_ρ −Cauchy sequence. For each $\varepsilon > 0$, we can say

$$\frac{1}{\rho_n}|\{k \le n: d(Z_k, Z_0) \ge \varepsilon\}| \le \frac{1}{\rho_n}\left|\left\{k \le n: d\left(Z_k, Z_{k'(n)}\right) \ge \frac{\varepsilon}{2}\right\}\right|$$
$$+\frac{1}{\rho_n}\left|\left\{k \le n: d\left(Z_{k'(n)}, Z_0\right) \ge \frac{\varepsilon}{2}\right\}\right|.$$

Hence, we get $Z_k \to Z_0\left(S_\rho(F)\right)$.

The proof to the contrary is obvious.

Theorem 2.4 Let (Z_k) be a fuzzy number sequence, $\rho = (\rho_n)$ is non-decreasing sequence for each $n \in \mathbb{Z}^+$ tending to ∞ such that $\limsup_n \frac{\rho_n}{n} < \infty$, $\Delta\rho_n = O(1)$ and $\Delta Z_n = Z_{n+1} - Z_n$ for each $n \in \mathbb{Z}^+$. If for each $n \in \mathbb{N}$, $\liminf\left(\frac{\rho_n}{n}\right) \ge 1$, then $S(F) \subset S_\rho(F)$.

Proof. Let's consider $Z_k \to Z_0(S(F))$, the following inequality leads to the proof, for every $\varepsilon > 0$

$$\frac{1}{n}|\{k \le n: d(Z_k, Z_0) \ge \varepsilon\}| = \frac{\rho_n}{n}\frac{1}{\rho_n}|\{k \le n: d(Z_k, Z_0) \ge \varepsilon\}|$$
$$\ge \frac{1}{\rho_n}|\{k \le n: d(Z_k, Z_0) \ge \varepsilon\}|.$$

Theorem 2.5. Let (Z_k) be a fuzzy number sequence, $\rho = (\rho_n)$ and $\varrho = (\varrho_n)$ be two sequences such that $\rho_n \le \varrho_n$ for all $n \in \mathbb{N}$. If $\liminf\left(\frac{\rho_n}{\varrho_n}\right) > 0$, then $S_\rho(F) \subset S_\varrho(F)$.

Proof. Suppose that $Z_k \to Z_0\left(S_\rho(F)\right)$, the following inequality leads to the proof, for every $\varepsilon > 0$

$$\frac{1}{\varrho_n}|\{k \le n: d(Z_k, Z_0) \ge \varepsilon\}| \le \frac{\rho_n}{\varrho_n}\frac{1}{\rho_n}|\{k \le n: d(Z_k, Z_0) \ge \varepsilon\}|.$$

Corollary 2.1 Let (Z_k) be a fuzzy number sequence, $\rho = (\rho_n)$ and $\varrho = (\varrho_n)$ be two sequences such that $\rho_n \le \varrho_n$ for all $n \in \mathbb{N}$. If $\liminf\left(\frac{\rho_n}{\varrho_n}\right) > 0$, then $S(F) \subset S_\rho(F) \subset S_\varrho(F)$.

Theorem 2.6. If $(Z_k) \to Z_0\left(w_\rho(F)\right)$, then $(Z_k) \to Z_0\left(S_\rho(F)\right)$.

Proof. Suppose that $(Z_k) \to Z_0\left(w_\rho(F)\right)$, for $\varepsilon > 0$, we can write

$$\frac{1}{\rho_n}\sum_{k=1}^{n} d(Z_k,Z_0) = \frac{1}{\rho_n}\left(\sum_{\substack{k=1\\ d(Z_k,Z_0)\geq\varepsilon}}^{n} d(Z_k,Z_0) + \sum_{\substack{k=1\\ d(Z_k,Z_0)<\varepsilon}}^{n} d(Z_k,Z_0)\right)$$
$$\geq \frac{1}{\rho_n}\sum_{\substack{k=1\\ d(Z_k,Z_0)\geq\varepsilon}}^{n} d(Z_k,Z_0)$$
$$\geq \varepsilon\frac{1}{\rho_n}\left|\{k\leq n: d(Z_k,Z_0)\geq\varepsilon\}\right|.$$

If we take the limit for $n\to\infty$, we have

$$\lim_{n\to\infty}\frac{1}{\rho_n}\left|\{k\leq n: d(Z_k,Z_0)\geq\varepsilon\}\right| = 0.$$

Thus, the desired outcome is obtained.

Corollary 2.2 Let (Z_k) be a fuzzy number sequence. If $(Z_k)\to Z_0(n\to\infty)$, then $(Z_k)\to Z_0\left(S_\rho(F)\right)$.

The opposite of the Theorem 2.6 and Corollary 2.2 aren't true, mostly. For example, let the $Z=(Z_k)$ sequence be as follows:

$$Z_k(z) = \begin{cases} \left.\begin{matrix} \frac{k}{k+2}z+\frac{2-2k}{k+2}, & z\in\left[\frac{2k-2}{k},3\right] \\ -\frac{k}{k+2}z+\frac{4k+2}{k+2}, & z\in\left[3,\frac{4k+2}{k}\right] \\ 0, & \text{otherwise} \end{matrix}\right\}, & \text{if} k=n^3\ (n=1,2,\ldots) \\ \left.\begin{matrix} z-2, & z\in[2,3] \\ -z+4, & z\in[3,4] \\ 0, & \text{otherwise} \end{matrix}\right\} := Z_0 & \text{if} k\neq n^3 \end{cases}$$

If we choose $(\rho_n)=n$,

$$\frac{1}{\rho_n}\left|\{k\leq n: d(Z_k,Z_0)\geq\varepsilon\}\right| = \frac{\sqrt[3]{n}}{n}\to 0(n\to\infty).$$

Moreover,

$$\frac{1}{\rho_n}\sum_{k=1}^{n} d(Z_k,Z_0) = \infty\ (n\to\infty)$$

so, Z_k is not $\left(w_\rho(F)\right)$ convergent Z_0.

Ethical statement

The author declares that this document does not require ethics committee approval or any special permission. Our study does not cause any harm to the environment.

Conflict of interest

The author declares no potential conflicts of interest related to this article's research, authorship, and publication.

References

[1] Fast, H., "Sur la convergence statistique", *Colloquium Math*., 2, 241-244, 1951.

[2] Schoenberg, I.J., "The Integrability of Certain Functions and Related Summability Methods", *Amer. Math. Monthly*, 66, 361-375, 1959.

[3] Connor, J., "A topological and functional analytic approach to statistical convergence", *Analysis of Divergence* , Birkhauser, Boston, 403-413, 1999.

[4] Fridy, J., "On statistical convergence", *Analysis* 5, 301-313, 1985.

[5] Šalát, T., "On statistically convergent sequences of real numbers", *Math. Slovaca*, 30, 139-150, 1980.

[6] Ercan, S., Altin, Y., Bektaş, Ç., "On lacunary weak statistical convergence of order $\boldsymbol{\alpha}$", *Communications in Statistics-Theory and Methods*, 49(7), 1653-1664, 2020 https://doi.org/10.1080/03610926.2018.1563185

[7] Zadeh, L. A., "Fuzzy sets", *Inform and Control*, 8, 338-353, 1965.

[8] Matloka, M., "Sequences of fuzzy numbers", *BUSEFAL*, 28, 28-37, 1986.

[9] Altinok, H., Çolak, R., Altin, Y., "On the class of $\boldsymbol{\lambda}$ −statistically convergent difference sequences of fuzzy numvers", *Soft Computing*, 16(6), 1029-1034, 2012 DOI: 10.1007/s00500-011-0800-6

[10] Aytar, S., Pehlivan, S., "Statistical convergence of sequences of fuzzy numbers and sequences of $\boldsymbol{\alpha}$ −cuts", *International Journal of General Systems* 37(2),231-237, 2008 DOI: 10.1080/03081070701251075

[11] UCakan, U., Altin, Y., "Some clases of statistically convergent sequences of fuzzy numbers generated by modulus function", *Iranian Journal of Fuzzy Systems*, 12(3), 47-55 , 2015.

[12] Çanak, İ., "On Tauberian theorems for Cesaro summability of sequences fuzzy numbers", *J.Intell. Fuzzy Syst*. 30, 2657-2661, 2016 DOI: 10.3233/IFS-131053

[13] Sezer, S.A., "Statistical harmonic summability of sequences of fuzzy numbers", *Soft Computing*, 27, 1933-1940, 2023 DOI: 10.1007/s00500-020-05151-9

[14] Tripathy, B.C., Baruah, A., "Lacunary statistically convergent and lacunary strongly convergent generalized difference sequences of fuzzy real numbers", *Kyungpook Math. Jour*. 50, 565-574, 2010.

[15] Nuray, F., Savaş, E., "Statistical convergence of sequences of fuzzy real numbers", *Math. Slovaca* 45(3), 269-273, 1995.

[16] Çakallı, H., "A variation on statistical ward continuity", *Bull. Malays. Math. Sci. Soc*. 40, 1701-1710, 2017. DOI: 10.1007/s40840-015-0195-0

[17] Kandemir, H.Ş., "On ρ-statistical convergence in topological groups*", Maltepe Journal of Mathematics*, 4(1), 9-14, 2022. doi:10.47087/mjm.1092559

[18] Aral, N.D., Kandemir, H.Ş., Et. M., "On $\boldsymbol{\rho}-$ Statistical convergence of sequences of Sets", *Conference Proceeding Science and Tecnology*, 3(1),156-159, 2020.

[19] Gumus, H., "Rho-statistical convergence of interval numbers", International Conference on Mathematics and Its Applications in Science and Engineering. 2022.

[20] Aral, N.D., Kandemir, H., & Et, M., "On $\rho-$statistical convergence of order α of sequences of function", *e-Journal of Analysis and Applied Mathematics*, 2022(1), 45-55, 2022.

[21] Aral, N.D., Kandemir, H., & Et, M., "On $\boldsymbol{\rho}$ $-$statistical convergence of double sequences of order $\boldsymbol{\alpha}$ in topological groups", *The Journal of Analysis,* 31, 3069–3078, 2023.

[22] Aral, N.D., Kandemir, H., & Et, M., "On $\boldsymbol{\rho}$ $-$statistical convergence of order $\boldsymbol{\alpha}$ of sequences of sets", *Miskolc Mathematical Notes*, 24(2), 569–578, 2023.

Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Research Article

MACHINE LEARNING APPROACH TOWARD TELEMARKETING ESTIMATION

Mehmet SALTİ[1] ***Evrim Ersin KANGAL***[2] ***Bilgin ZENGIN***[3]

[1]Department of Business Information Management, Graduate School of Social Sciences, Mersin University, Mersin, Türkiye
[2]Computer Technology and Information Systems, School of Applied Technology and Management of Erdemli, Mersin University, Mersin, Türkiye
[3]Department of Electrical and Electronic Engineering, Faculty of Engineering, Munzur University, Tunceli, Türkiye
Corresponding author;evrimersin@gmail.com

Abstract: *Machine learning empowers us to extract insights from large datasets beyond human capacity. It involves training computers to identify patterns within data, enabling them to glean valuable information and apply it to novel tasks. This study focuses on analyzing a specific telemarketing dataset using various machine learning algorithms to determine if accurate predictions can be made to support company decision-making. The findings highlight that customer "Age" and "Product ID" are the primary factors influencing "Sales" numbers, indicating their significance in the predictive model.*

Keywords*: Data mining, machine learning, kNN, telemarketing.*

Received: January 28, 2024 *Accepted: April 3, 2024*

1. Introduction

Marketing campaigns are constructed via a typical strategy to enhance business capacity, which is a measure of the amount of work that may be overcome by a company within a fixed amount of time [1]. Contacting different types of customers to meet specific targets, companies use direct strategies while communicating with them. On the way to this goal, companies prefer to centralize their remote interactions with customers in a contact center to facilitate the operational management of campaigns. Such contact centers communicate with potential customers through various channels: telephone (landline or mobile) is the most widely used tool in this context. Marketing carried out through a contact center is called telemarketing in literature because of its remoteness feature [2]. Depending on which factor (customer or contact center) triggers, people are divided into two groups inbound and outbound. It is worth emphasizing here that each situation presents different challenges (for instance, outgoing calls are often seen as annoying). Technology requires us to rethink marketing methods and aim to increase the value of a customer by addressing information and demands from different perspectives, thus it enables us to establish longer and tighter relationships in line with business demands [3].
Developing technology has led to an exponential increase in the global data volume. One of the sources from which intensive scientific data is produced today is the large hadron collider (LHC) at the European Organization for Nuclear Research (CERN) [4]. The LHC produces a continuous data stream of 40 Tb/s [5]. Such datasets are called “Big Data” [6] in literature. As expected, the number of recorded features of samples also increases due to an aggressive enlargement in the data volume. On the other hand, the

increase in the recording area causes “the curse of dimensionality” [7]. Marron and Maine [8] thought that this problem would be overcome with the data compression method, but in the following years, it was revealed that the data compression causes some information to be lost [9-11]. Data analysis through artificial intelligence (AI) has gained considerable popularity in recent years. AI is generally taken into account in literature to identify the artificial creation of human brains that can learn, plan, perceive, or process [12] and it is basically the development of computer systems capable of performing tasks (for instance visual perception, speech recognition, decision-making, and language translation) that need human intelligence [13]. ML, Neural Networks, and Deep Learning (DL) are sometimes used interchangeably, but there are some nuances among them. All of these concepts are actually sub-fields of the AI: the DL is a sub-branch of Neural Networks while a neural network approach is a sub-branch of the ML. The AI approach has noteworthy potential to increase the level of our knowledge in diverse fields such as finance [14], business management [15], meteorology [16], neuroscience [17], demoscopic analysis [18], quantum chemistry [19], spacecraft engineering [20], astrophysics [21-24], cosmology [25-27], etc.

The ML algorithms are divided into two main parts: the supervised architectures and the unsupervised ones [28,29]. The supervised learning mechanisms try to establish a relationship, which is organized by the classification or regression algorithms, between the input data and the corresponding output. On the contrary, in the case of unsupervised learning, there is no use of any training set and data, which are divided into different branches according to similarity criteria involving data features. In addition, an ordinary decision support system (DSS) focuses on information technology to support the managerial decision-making phase and there are various DSS types, such as personal and intelligent DSSs [1]. A personal DSS takes an effective role in relatively small-scale situations (for example, a manager's decision-making task), while an intelligent DSS uses AI algorithms to support decisions [30]. Business intelligence (BI) is another important DSS concept. The BI is generally accepted as an umbrella term that supports decision-making tasks by making use of business data and consists of information technologies such as data warehouses and data mining [31]. Classification algorithms used commonly in the field can be listed as follows [32-36]: the Decision Trees (DTs), Linear Regression (LiR), Logistic Regression (LoR), Support Vector Machines (SVMs), Naive Bayes Classifier (NBC), Neural Networks, Instance-based Learning (IBL) and the k-Nearest Neighbors (kNN).

Telemarketing is a sort of direct marketing where salespersons communicate with prospective customers to acquire new customers, vend their products, and provide their services over the phone [2]. Since only the direct marketing dataset helps us to reach the database of prospective customers, it is very significant for a company to estimate successfully the group of potential clients with the highest prospect to admit the sales and/or offer according to their individual characteristics or attitude while shopping [37]. Therefore, many companies have recently started to focus on data mining methods for customer classification while many scientists have taken diverse ML applications into account for the different telemarketing databases. Tekouabou et al. [38] introduced a new data modeling approach to optimize the estimation of telemarketing target calls for selling bank long-term deposits. Keles and Keles [39] are interested in constructing an Intelligent Bank Market Management System based on the ML approaches to operate marketing campaigns efficiently. Kocoglu and Esnaf [40] developed a model to classify the success of telemarketing with various ML approaches, such as the NBC, C5.0, Extreme Learning Machine (ELM), and the DL. Considering a telemarketing dataset, Halim et al. [41] concluded that the Neural Networks approach, which is supported with data cleaning algorithms such as the Missing Common and the Tomek Links, indicates a better conclusion compared to the Ignore Missing mechanism. Shashidhara et al. [42] have investigated the most appropriate model for the analysis of marketing data in the banking sector with ML architectures.

In the present study, the focus is aimed at the wrapper-based kNN algorithm for conducting data analysis in a telemarketing case. The outline of this paper is as follows. In the second section, the methodology and the selected data pool are introduced. In the third section, the ML architecture is constructed, and the selected performance testing method is introduced. Next, in the fourth section, attention is given to preprocessing and visualization of the dataset to decipher some hidden features of the case. The constructed wrapper-based kNN algorithm is tuned in the fifth section. In the sixth section, the architecture is run. The final section is devoted to final remarks. It is noted that Python Anaconda 3.7 version is used for coding all algorithms and as an interpreter. Furthermore, the Pandas, Seaborne, Scikit, Numpy, and Matplotlib libraries are taken into account extensively in the phase of architecture implementation.

2. Methodology and Dataset

Increasing campaign costs and low marketing-response rates drive companies to model customer behavior to obtain successful sales statistics, but this requirement often forces them to work with complex methods. It is known that predicting a customer's reaction to marketing may provide a significant advantage to a company before the campaign. Motivated by this situation, the problem of telemarketing success classification is discussed within the scope of our study. It is carried out the investigation in accordance with the well-known methodology CRoss Industry Standard Process for Data Mining (CRISP-DM), which has six sequential phases [43,44]:

- Understanding the problem - What the need of business is?
- Deciphering the data - What data do we have? What do we need?
- Data preparation - Is it clean? How do we organize it before modeling?
- Developing the appropriate techniques - Which models can we apply?
- Evaluating performance of the approach - Which algorithm best meets the goals?
- Application - How do stakeholders achieve the goals?

In this study, we focus on an anonymous telemarketing case, which is taken from the Kaggle platform and includes high-quality datasets[45]. The selected data pool is represented by a table that includes 12 feature columns and 100000 rows of customer information. A part of the aforementioned dataset is presented in Table 1 as an example.

Table 1. A sample is extracted from the dataset. [45]

Call ID	Sales	Agent ID	Age	Product ID	Time Zone	Phone Code	First Name	Last Name	Area Code	Gender	Call Count
9545434	False	5265	42	147	2	37	Jk	Jk	2302	Male	1
9211206	False	5226	74	146	2	37	Em	Sh	1501	Male	10
8873010	False	4452	35	144	2	37	BI	MI	1550	Male	9
9852034	False	5461	40	149	2	37	WT	LI	1401	Male	6
9416548	False	5298	26	147	2	37	LA	LA	125	Female	12
10189322	False	5139	33	150	2	37	Me	Is	4091	Female	1
10277850	False	4828	33	151	2	37	So	Ts	3880	Female	3
9105514	False	5292	70	145	2	37	Dy	Ma	432	Female	9
8663012	False	5044	41	143	2	37	No	Mi	4360	Female	9
10216124	True	4912	72	150	2	37	TA	MS	1983	Male	1

It would be appropriate to emphasize here that extremely serious problems may arise for the ML algorithms at the distribution classification phase. If the selected architecture is trained via the samples identified with insufficient label information, the algorithm is likely to make useless predictions. It is seen that attributes in Table 1 are represented via different data types (please check Table 2).

Table 2. Attribute types in the dataset

Attribute	Type	Attribute	Type
Call ID	Integer	Phone Code	Integer
Sales	Categorical	First Name	Categorical
Agent ID	Integer	Last Name	Categorical
Age	Integer	Area Code	Integer
Product ID	Integer	Gender	Categorical
Time Zone	Integer	Call Count	Integer

In this statistical analysis, the attribute "Sales" is selected as the target feature while all remaining attributes are taken into account as arguments. It is obvious that the “Time Zone”, Phone Code”, “First Name”, “Last Name” and “Call ID” attributes do not significantly affect the attribute “Sales”. Subsequently, in the entire table, some customers have deficiencies in some of their information and we conclude that missing information constitutes approximately 3% of the entire data pool. Accordingly, the data cleaning process to be performed in Table 1 enables faster and more consistent completion of our analyzes. Our dataset now contains 223 Agent IDs, 10 product variants, and 97205 customers. On the other hand, it is mainly to discuss whether an ML pattern can be constructed among the independent variables and the selected target parameter, thus the categorical information in Table 1 needs to be transformed into binary values with the help of dummy variables. Once the label encoder tool in Python has been applied, restructuring the dataset becomes a much more straightforward task. The reader may check Table 3 to see the assumptions we used for the numerical conversion of the categorical attributes.

Table 3. Numerical conversion of the categorical information

Feature	Gender (Male)	Gender (Female)	Gender (Non-binary)	Sales (True)	Sales (False)
Value	1	0	2	1	0

Moreover, category imbalance is another situation that needs to be carefully addressed in data science research. Most of the classification algorithms perform analysis by assuming that the presented raw dataset is balanced, but this is not the case often. Some of the categories (classes) defined in the repository may have very few elements while others may be represented by too many elements. In such cases, the selected classification algorithm yields serious issues and makes erroneous predictions, because it cannot be trained well enough for dataset elements with low characteristic information. In classification approaches, the main goal is to maximize the beneficial prediction rate, thus necessitating the balancing of attribute distributions through a suitable technique. For this purpose, in FIG. 1, it is discussed as a distribution of the results given in the "Sales" column of Table 1 and see that the number of “False (0)” cases is much higher than the “True (1)” ones. Such a situation may affect negatively the performance of our ML approach; Hence, the need to synthesize new "True (1)" cases via a suitable Python package is recognized to address this distribution problem.

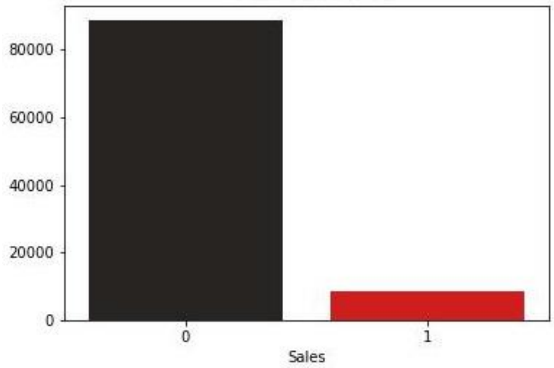

Figure 1. The number of “False (0)” and “True (1)” sales

3. Architecture

Nearest neighbor-based algorithms produce very successful results in the classification research and they treat the information obtained from the dataset in the form of specific situations or experiences. As they rely on efficient matching methods to retrieve stored information, they can be applied to new situations. The main purpose of such methods is to determine the nearest neighbors of a new data point and then place it in the correct class. Therefore, for such approaches, it is of great importance to calculate the shortest distance between two data points. We assume here that the function $d(x, y)$ basically identifies the distance between two different feature vectors $A = (x_1, x_2, \dots, x_k)$ and $B = (y_1, y_2, \dots, y_k)$ as a non-negative real number. A distance function is assumed to be a metric if it obeys a certain number of features that consist of the following conditions: positivity (or non-negativity), the identity of the inseparable, symmetry, and triangle inequality (the reader may check Refs. [46,47] for detailed information).

Example-based ML algorithms are computationally simple and generally recognized as different types of human learning [48]. The Instance-based Learning (IBL) is a simple nearest neighbor-based ML algorithm [49]. Here, the closest data point to the sample to be classified is searched according to the Euclidean formalism and then the sample is assigned to the class of that data point. On the other hand, the class of the unknown sample is determined by the majority of its neighbors. It is seen from FIG. 2 that, in the K_1 region, new data points are estimated as members of group B since the purple square data points are the dominant. Next, the majority of orange hexagonal data points in the K_2 region indicate that the new data points are members of group A.

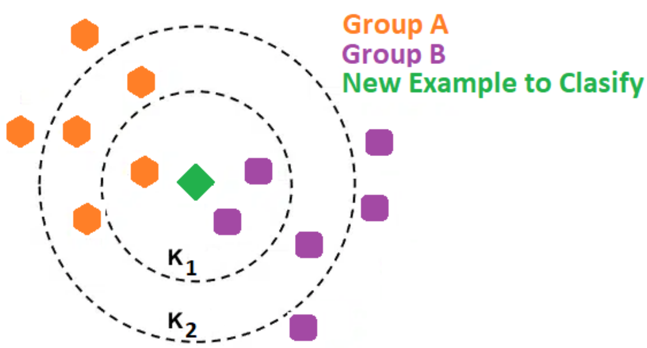

Figure 2. A representative visualization of the nearest neighbor-based classification

The related distance measurement metric used by the IBL algorithm is expressed as follows [50]:

$$d_{Euclidian}(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}, \quad (1)$$

where x_i and y_i represent the i-th values of x and y samples, respectively.

The learning process is based on the selected training data also in the kNN algorithm, essentially the kNN mechanism is a more general form of the IBL approach. Here, k represents the number of nearest neighbors. Unlike the IBL approach, the distance between the two closest data points in the kNN mechanism can be measured by various methods. From this point of view, we complete our ML analysis according to kNN algorithms designed according to the 8 most used distance measurement formalisms, including the Euclidean metric. We consider the Euclidean, CityBlock (also known as the Manhattan, Taxicab, Rectilinear, or the L_1-norm), Chebyshev, Correlation, Canberra, Dice, and the Cosine measurements during the patterning phase in our research. Here, we have [50,51]

$$d_{CityBlock}(x,y) = \sum_{i=1}^{n}|x_i - y_i|, \quad (2)$$

$$d_{Chebyshev}(x,y) = \max_i |x_i - y_i|, \quad (3)$$

$$d_{Correlation}(x,y) = \frac{1}{2}\left(1 - \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}\right), \quad (4)$$

$$d_{Canberra}(x,y) = \sum_{i=1}^{n}\frac{|x_i-y_i|}{|x_i|+|y_i|}, \quad (5)$$

$$d_{Dice}(x,y) = 1 - \frac{2\sum_{i=1}^{n}x_i y_i}{\sum_{i=1}^{n}{x_i}^2+\sum_{i=1}^{n}{y_i}^2}, \quad (6)$$

$$d_{Cosine}(x,y) = 1 - \frac{\sum_{i=1}^{n}x_i y_i}{\sqrt{\sum_{i=1}^{n}{x_i}^2}\sqrt{\sum_{i=1}^{n}{y_i}^2}}, \quad (7)$$

where n represents a dimension of the feature space and the upper bar indicates the mean value of the relevant quantity.

The increase in the number of features not only negatively affects the algorithm in the decision phase and but also prolongs the compilation time of the mechanism. This issue is called “the curse of dimensionality” in literature [52-54] and can be removed by making use of the Feature Selection (FeSe) approach, which is one of the widely used solutions in the field. The essence of such methods is based on the determination of attributes that effectively influence the decision-support system. Consequently, unnecessary features are eliminated from the selected dataset and the ML pattern required for the desired performance is now formed more easily [55]. In this investigation, the wrapper-based supervised FeSe approach is considered. According to this method, feature subsets are created first, and then the results of each selected subset are compared with random selections. For this purpose, four different wrapper types are used to achieve the best results: the Sequential Backward Selection (SBS), the Sequential Forward Selection (SFS), the Sequential Backward Floating Selection (SBFS), and the Sequential Forward Floating Selection (SFFS). The algorithm creates a universal set by removing subsets one by one in the SBS approach while a universal attribute pool is created by adding the attributes one by one to the dataset in the SFS mechanism [56,57]. The main point here is that the features covered are selected randomly from the feature space. On the other hand, both the SBFS and SFFS methods were born as alternative models to the Plus-L-Take-Away-R (or the Plus-L-Minus-R) algorithm in essence [58]. Since there is no a theoretical approach to determine the values of the parameters L and R, the performance percentage of an ML algorithm becomes directly dependent on these auxiliary parameters, which means

the choice of L and R is the key to success of this algorithm [59]. In order to eliminate this dependence, the floating (or moving) L and R values are usually created instead of the fixed ones [60].
The workflow of our architecture is schematically illustrated in FIG. 3 and one can check here the main procedure of the selected ML algorithm.

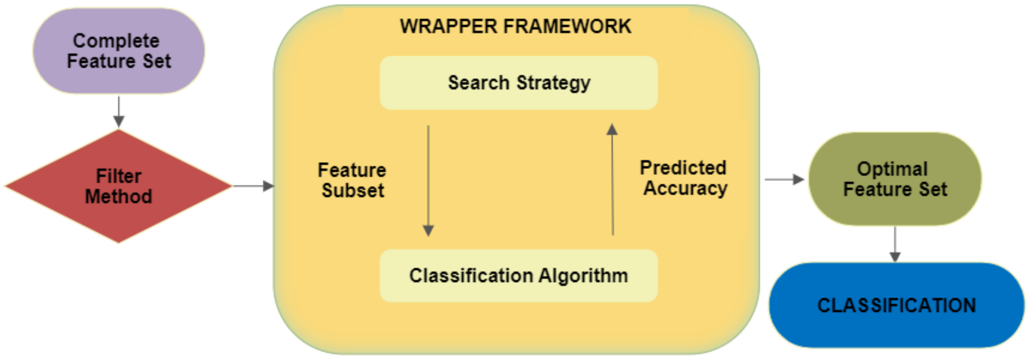

Figure 3. Flow chart of the wrapper-based kNN mechanism

There are 8 different distance measurement metrics, 4 types of wrappers and different k values and it is aimed to build a successful kNN algorithm. For this reason, there is a need for a performance test method that can allow us to compare different algorithm combinations.

		Estimated Data			
		Positive (1)	**Negative (0)**		**Totals**
Actual Data	**Positive (1)**	***TP***	***FN***	Sensitivity (Positive Recall) $\frac{TP}{TP+FN}$	Actual Positives $P = TP + FN$
	Negative (0)	***FP***	***TN***	Specificity (Negative Recall) $\frac{TN}{FP+TN}$	Actual Negatives $N = FP + TN$
		Positive Precision $\frac{TP}{TP+FP}$	Negative Precision $\frac{TN}{TN+FN}$	Accuracy $\frac{TP+TN}{TP+FN+FP+TN}$	
	Totals	Estimated Positives $P = TP + FN$	Estimated Negatives $N = FN + TN$		

Figure 4. A general representation of confusion matrix

Performance evaluation of an ML processes can be made by means of a number of metric criteria [61]. The confusion matrix is among the well-known criteria and it is structured in a way that we can obtain True and False values as a result of classification. In FIG. 4, we present a general representation of confusion matrix. Here, rows correspond to actual data while columns imply predicted data. Also, the abbreviations TP, FN, FP and TN mean respectively True Positive (we have actual positivity in the data, which has been estimated correctly as positive by the selected model), False Negative (there is an actual negativity in the data, which has been predicted accurately as negative by the algorithm), False Positive (we have actual negativity in data, but the selected approach has predicted it as positive) and True Negative (there is an actual positivity in data, but the model has estimated it as negative). In the light of these definitions, various mathematical metrics such as the ACC (Automatically determining the Cluster Centers) [62], RAC (Relative Angle Correction) [63] and F-Measure (or F1-score) [64] can also be used to test the performance of an ML approach.

4. Preprocessing and Visualization of Data

An unbalanced dataset may be a serious source of inconsistency when trying to use the selected ML algorithm in the decision-support phase. There are two main ways to eliminate this issue: under-sampling and over-sampling. On this purpose, one can use the SMOTE (Synthetic Minority Oversampling Technique), NCL (Neighborhood Cleaning Rule), OSS (One-Sided Selection) and the BootsOS (Bootstrap-based Over Sampling) approach [65-69]. In FIG. 5, it is illustrated as the process of creating a synthetic dataset via the under-sampling and over-sampling approaches. Here, the red group (A) is the minority class while the black one (B) is the majority class.

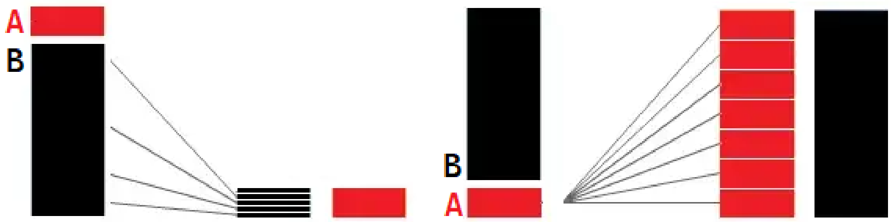

Figure 5. Under-sampling (left) and over-sampling (right) mechanisms

Based on the situation shown in FIG. 1, Sufficient synthetic "True (1)" cases for the attribute "Sales" are generated with the help of SMOTE. After this task, FIG. 6 is obtained, which indicates that the number of “True (1)” and “False (0)” sales are now balanced.

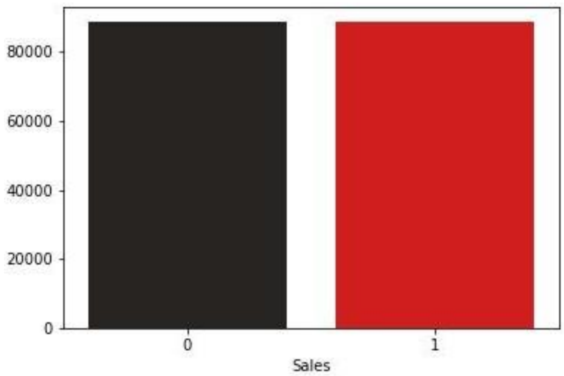

Figure 6. The number of “False Sales” (Black) and “True Sales” (Red) after the SMOTE

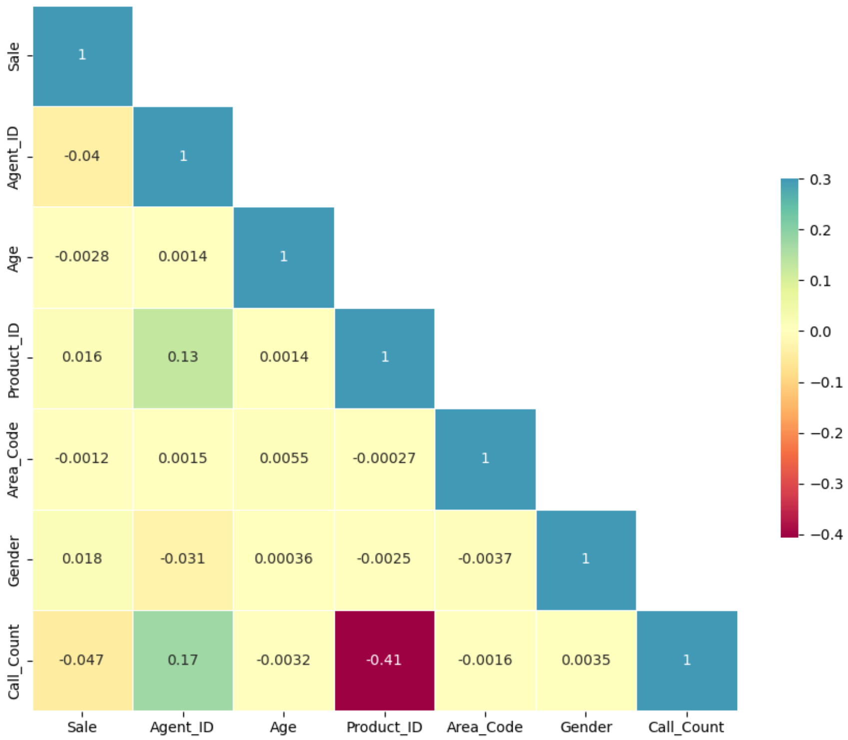

Figure 7. Correlation analysis of the attributes

Another important point that can negatively affect a ML analysis is the correlation between attributes. A correlation table includes columns and rows that represent variables of the selected dataset. A negative correlation between any two features is indicated when the corresponding correlation value falls within the range of [-1, -0.5]. On the other hand, there is a positive correlation between any two attributes if the correlation coefficient is in the range of [0.5, 1]. The main diagonal line that includes 1.00s implies that each variable always perfectly correlated with itself and other correlation values indicate no-correlation cases. FIG. 7 is created to reveal the positive and negative correlations among the attributes of our dataset clearly. It is seen that there is no issue that may adversely affect the performance of our ML analysis, since there is no connectivity problem among the attributes.

Over-sampling process may have changed other properties of the repository as well, hence controlling these changes can allow us to access additional information about the selected tele-marketing case. In FIG. 8, we check the gender distribution in the dataset. Subsequently, in FIG. 9, we aim to discuss how the number of customers is distributed by the attribute “Age” before and after the SMOTE. In our investigation, the SMOTE algorithm focused on the available raw data while synthesizing the attribute labeled "True Sales", thus an increase in the number of female customers was observed. This indicates that the positive marketing situations in the raw data set are mostly observed in the female customer profile.

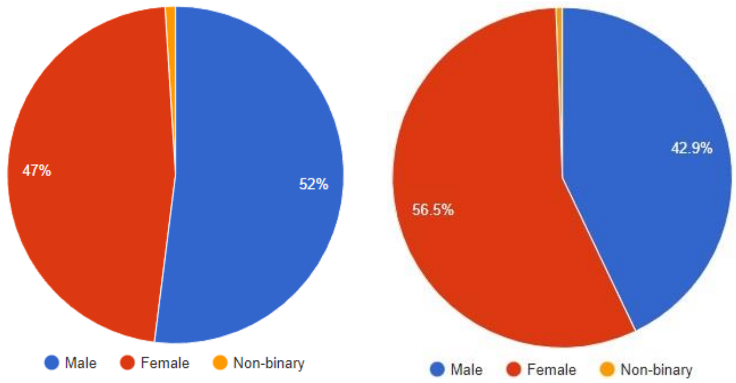

Figure 8. “Gender” distribution before (left) and after (right) the SMOTE

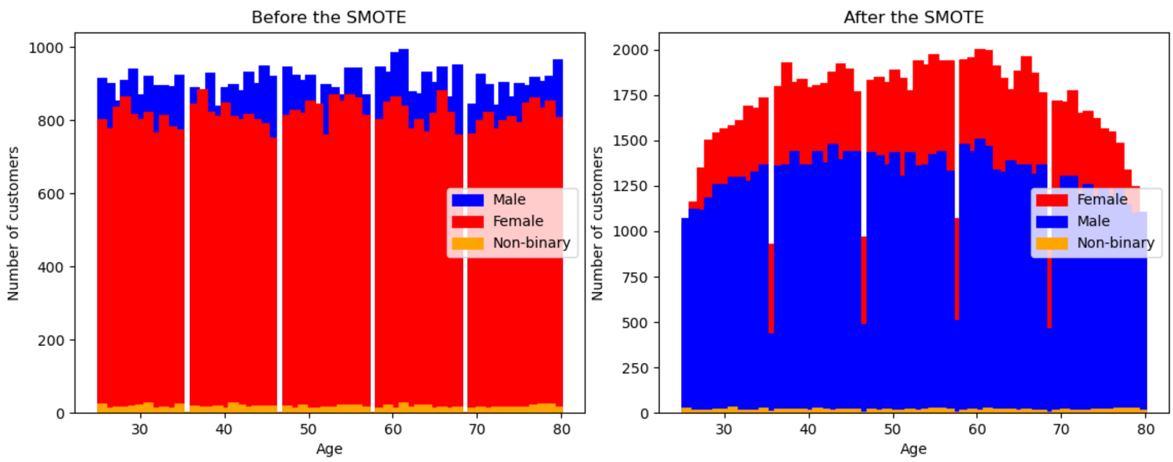

Figure 9. Distribution of the number of customers according to their ages

At FIGs. 10 and 11, it is discussed the distribution of the number of “Sales” according to the attributes “Agent ID” and “Product ID”. It is understood that some agents put more effort in reaching customers and some products are marketed more than others. In addition, it is seen that the SMOTE algorithm preserves characteristic features of the raw data. This is important because it tells us that the estimated results obtained as a result of the ML application can be trusted.

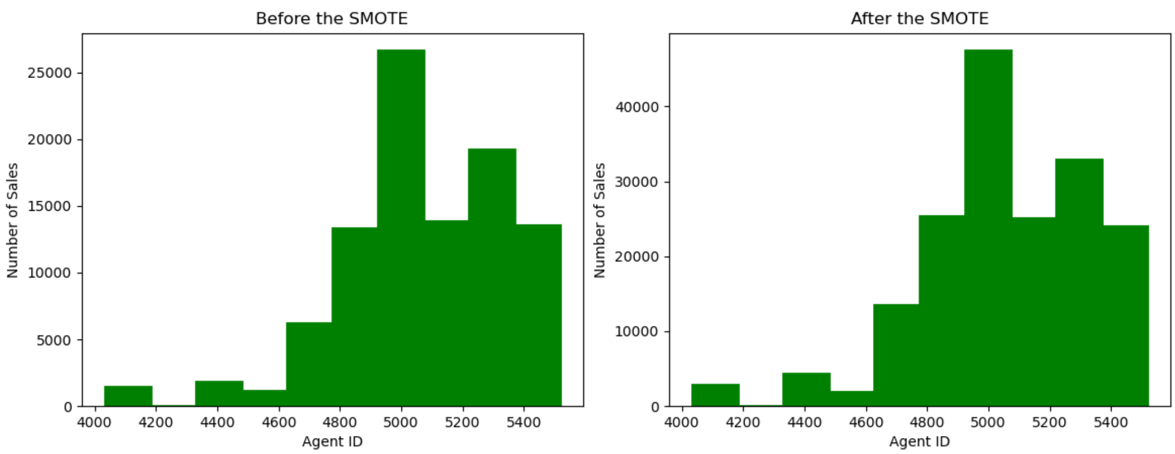

Figure 10. Number of “Sales” vs “Agent ID”

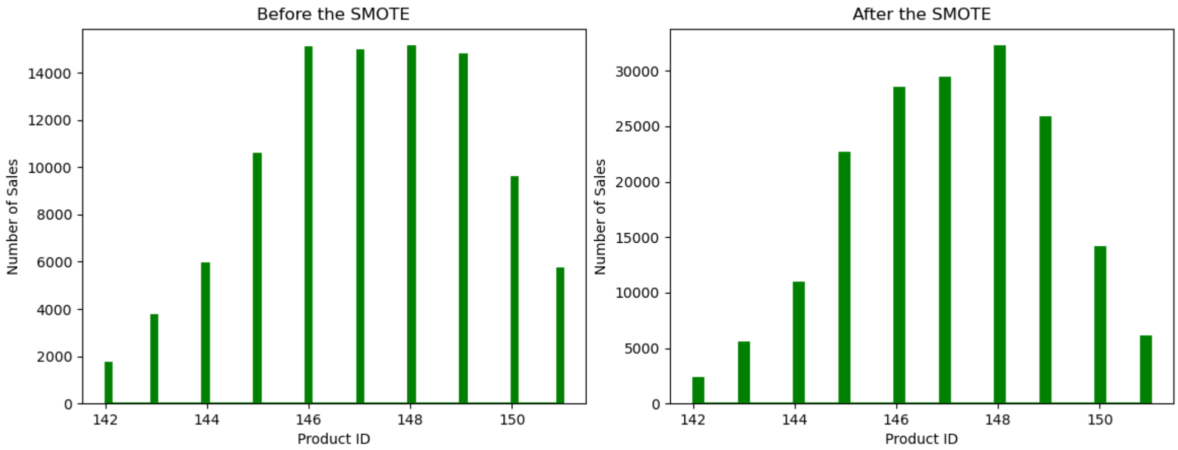

Figure 11. Number of "Sales" vs "Product ID"

5. Tune the Algorithm

A hyperparameter (the prefix "hyper" indicates here that such a parameter is top-level quantity) is a vital ML part that is explicitly defined by the user to control the learning process. For example, batch size, number of epochs, pooling size, learning rates, number of hidden layers, and the number of neighbors are some of the well-known hyperparameters [70-72].
In this part of the study, we do not only aim to determine the most reasonable value of the hyperparameter *k*, which plays an important role in the wrapper-based kNN algorithm, but also want to choose the appropriate "distance function" and "wrapper" formulations. In other words, the main idea is to figure out the best combination of algorithm parts. The general method is to use some of the available data as a tuning group in different algorithm combinations. Consequently, the various kNN algorithms were run by considering the train-test split ratio of 70-30, different values of *k*, seven kinds of distance function and four different wrapper types.

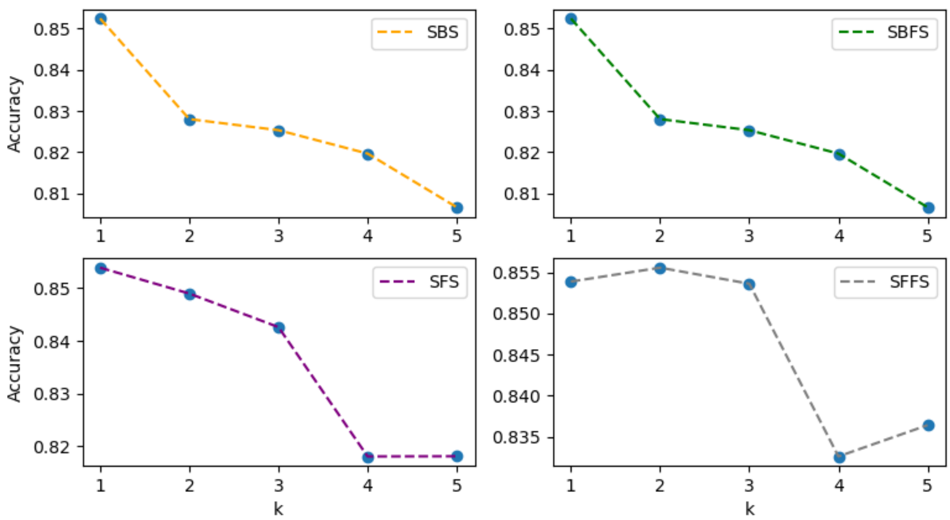

Figure 12. Performance analysis of the CityBlock measurement

Table 4. Accuracies obtained in the CityBlock analysis

k	SBS	SBFS	SFS	SFFS
1	0.852453	0.852453	0.853901	0.853901
2	0.828050	0.828050	0.848994	0.855593
3	0.825362	0.825362	0.842621	0.853638
4	0.819684	0.819684	0.818067	0.832581
5	0.806655	0.806655	0.818105	0.836398

Table 5. Tuning results according to the highest accuracy values

Distance Type	*k*	Wrapper	Accuracy (%)	Process Time (≈)
CityBlock	2	SFFS	85.6	13 minutes
Chebyshev	2	SFFS	85.1	1 hour
Euclidean	2	SFFS	85.3	45 minutes
Correlation	1	SFFS	87.5	7 hours
Canberra	2	SBS, SBFS, SFFS	91.6	9 hours
Dice	1	SFS, SFFS	62.6	11 hours
	5	SBS, SBFS, SFS, SFFS		
Cosine	2	SFS, SFFS	87.9	10 hours

For the CityBlock distance measurement function, we reach the results presented in FIG. 12 and Table 4. Subsequently, the same procedure is followed for the Chebyshev, Euclidean, Correlation, Canberra, Dice and the Cosine type distance measurements. It can be summarized the obtained noteworthy results in Table 5. In conclusion, the most appropriate combination of the architecture components and the best value of the hyper parameter *k* are now clearly revealed. The most significant differences among the performance of various distance measurement expressions are the maximum accuracy and the process time. Although the Correlation, Canberra and the Cosine formulations offer relatively higher prediction accuracy, it is understood that they require a remarkably long analysis time and a powerful set of computer hardware. Moreover, all types of kNN architecture give their best results while working with the SFFS type wrapper.

6. Run the kNN architectures

The ML analysis is now sought to be deepened to explore the implications of different kNN architectures on the attribute "Sales" in our telemarketing case. To achieve this goal, it is assumed that the train-test split ratio of the selected dataset is 70-30 again, and the kNN architectures consist of the SFFS type wrapper.

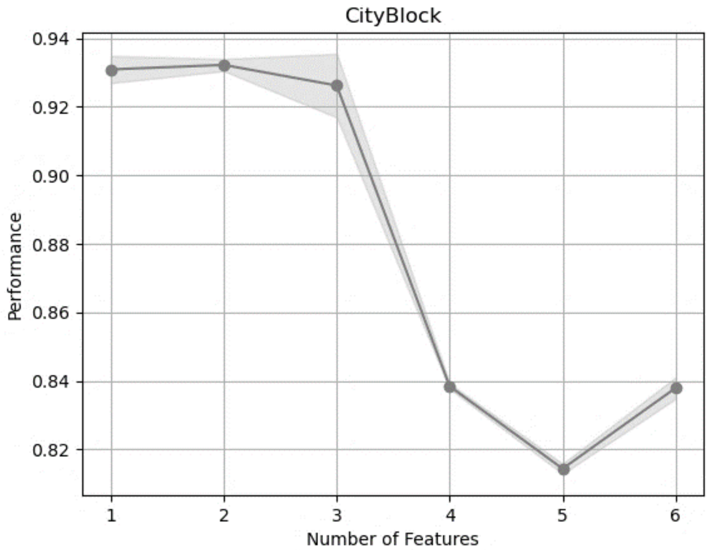

Figure 13. Performance analysis of the CityBlock_SFFS_kNN_vk2 algorithm

First, the performance of the CityBlock measurement based SFFS-kNN algorithm (we name this architecture CityBlock_SFFS_kNN_vk2) is obtained as illustrated in Figure 13. Subsequently, we give a detailed report of this 1-minute analysis in Table 6. It is basically concluded that the success rate of the CityBlock-based algorithm decreases as the number of features affecting the decision increases. According to the CityBlock_SFFS_kNN_vk2 architecture, the most important features affecting the attribute "Sales" are the "Age" of customer and the "Product ID".

Table 6. Detailed performance analysis of the CityBlock_SFFS_kNN_vk2 algorithm

Number of Features	**1**	**2**	**3**	**4**	**5**	**6**
Accuracy (%)	**93.1**	**93.2**	**92.6**	**83.8**	**81.4**	**83.8**
Agent ID				X	X	X
Age	X	X				X
Product ID		X	X	X	X	X
Area Code					X	X
Gender			X	X	X	X
Call Count			X	X	X	X

Next, for the CityBlock_SFFS_kNN_vk2 algorithm, the components of the confusion matrix are obtained as given below

$$C_{CityBlock} = \begin{pmatrix} TN & FN \\ TP & FP \end{pmatrix} = \begin{pmatrix} 26383 & 269 \\ 3249 & 23289 \end{pmatrix}. \quad (8)$$

Consequently, the performance values are calculated as presented in Table 7.

Table 7. Performance report of the CityBlock_SFFS_kNN_vk2 model

Weighted Mode (0,1)	Precision	Recall	F1-score	Support
0	0.89	0.99	0.94	26652
1	0.99	0.88	0.93	26538
Total				
Accuracy			0.93	53190
Macro Average	0.94	0.93	0.93	53190
Weighted Average	0.94	0.93	0.93	53190

The same evaluation steps are followed respectively also for the Chebyshev, Euclidean, Correlation, Canberra, Dice and the Cosine distance models. In conclusion, the main results are summarized in Table 8.

Table 8. Comparing all types of the kNN architecture

Architecture	Accuracy Value	Performance Rate	Effective Attributes	Process Time (≈)
CityBlock_SFFS_kNN_vk2	0.93	0.9322	Age, Product ID	1 minute
ChebyShev_SFFS_kNN_vk2	0.93	0.9327	Age, Product ID	1 minute
Euclidean_SFFS_kNN_vk2	0.93	0.9336	Age, Product ID	1 minute
Correlation_SFFS_kNN_vk1	0.90	0.9032	Age, Gender, Product ID	47 minutes
Canberra_SFFS_kNN_vk2	0.94	0.9329	Age, Gender	40 minutes
Dice_SFFS_kNN_vk1	0.63	0.5983	Gender, Call Count	75 minutes
Dice_SFFS_kNN_vk5	0.63	0.6225	Product ID, Gender	65 minutes
Cosine_SFFS_kNN_vk2	0.94	0.9322	Product ID, Gender, Call Count	70 minutes

7. Final Remarks

Human expectations and technological developments take vital roles in increasing the significance of data and knowledge nowadays. In this direction, the data gathered from various sources are kept in different data storage environments. Consequently, various approaches are introduced in literature to process the data stacks, which are increasing in size and changing in structure. The ML approach, which is about focusing on mathematical equations representing real-world scenarios, is among these methods. The key tasks of ML are regression, classification, clustering, transcription, machine translation, anomaly detection, synthesis & sampling, estimation of probability density, prediction of probability mass function, similarity matching, so-occurrence grouping, causal modeling and link profiling. On the other hand, the main development phases that have been used to overcome the aforementioned ML tasks are exploratory data analysis, data preprocessing, feature engineering (feature creation/extraction, feature selection, feature dimensionality reduction), architecture selection, training algorithms, testing & matching, approach monitoring and model retraining. Within the scope of this research, the analysis of a telemarketing dataset of company X through a classification-capable ML architecture is intended, and its methodology is illustrated in Figure 14.

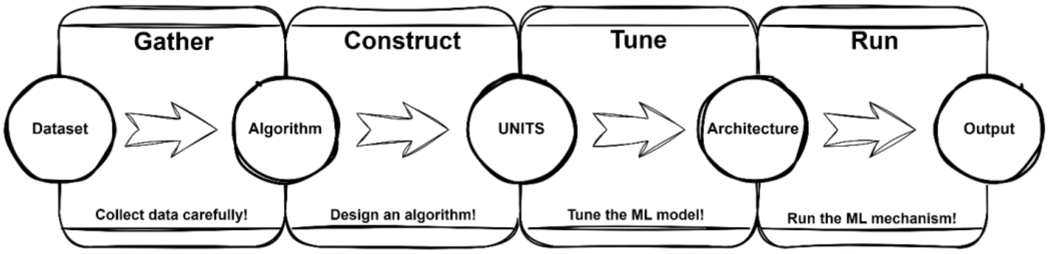

Figure 14. An illustration for our methodology

An ordinary ML architecture consists of various subunits and critical hyperparameters, therefore choosing the most appropriate combination of components is a critical step for the success of the algorithm. Unfortunately, no algorithm can learn the most appropriate subunits and determine the best value of hyperparameters. In the first phase of our research, the dataset has been cleaned, arranged and balanced. These tasks can be done with a series of commands including various packages of the Python. The second important phase in our research was the design of the most appropriate ML architecture. This task generally dominates a significant portion of an ML research in terms of research time spent. Various subunits and different hyperparameter values have been considered in our research, and then some suitable kNN architectures have been constructed to analyze the selected telemarketing case. In the last step of our investigation, we have run these suitable forms of the kNN architecture. A remarkable conclusion has emerged from the results presented in Table 8: the “Age” of customer and “Product ID” are the most important attributes that affect the “Sales” column in Table 1. If a preference is given to an analysis completed quickly, the classification success rate of the architecture relatively decreases. Conversely, if the aim is to achieve the most successful forecasting results, time must be sacrificed.

Figure 15. Best combination of the kNN architecture in our investigation

As a matter of fact, the most appropriate architectural design clearly comes to the fore. According to the highest accuracy rate and performance value, the best kNN algorithm should be as in FIG. 15. The architecture above shows that the most important factors affecting the "Sales" attribute are the "Age" and "Gender" information. So, considering the whole of our research, it would be appropriate to add "Product ID" information next to these two effective attributes.

Ethical statement

The data is sourced from an open-access database, so there is no need for an ethics committee’s evaluation.

Acknowledgment

We are grateful to the anonymous referees for substantial suggestions that improved the quality of our study. Also, we would like to thank Assoc. Professor Mustafa SALTI for their valuable financial support.

Conflict of interest

The authors declare no competing interests.

Authors' Contributions

M. Salti: Software, Formal analysis, Data curation. E.E. Kangal: Supervision, Conceptualization, Methodology, Writing-original draft, Writing - review & editing B. Zengin: Validation, Investigation.

References

[1] Moro, S., Cortez, P., Rita, P., "A data-driven approach to predict the success of bank telemarketing", *Decision Support Systems*, 62, 22-31, 2014.

[2] Kotler, P., Keller, K.L., "*Framework for Marketing Management (6th edition)*", Pearson, London-UK, 2015.

[3] Rust, R.T., Moorman, C., Bhalla, G., "*Rethinking Marketing*", Harvard Business Review, 1, 1, 2010.

[4] The reader can find detailed information about the European Organization for Nuclear Research (CERN) at https://www.home.cern.

[5] R. Krawczyk, Colombo, T., Neufeld, N., Pisani, F., Valat, S., "Ethernet for high-throughput computing at CERN", *IEEE Transactions on Parallel and Distributed Systems*, 33, 3640-3650, 2022.

[6] Buhl, H.U., Roglinger, M., Moser, F., Heidemann, J., "Big Data", *Business & Information Systems Engineering*, 5, 65-69, 2013.

[7] Verleysen, M., Francois, D., "The Curse of Dimensionality in Data Mining and Time Series Prediction", *Computational Intelligence and Bioinspired Systems*, 3512, 758-770, 2005.

[8] Marron, B. A., de Maine, P. A. D., "Communications of the ACM", *Communications of the ACM*, 10, 711-715, 1967.

[9] Heavens, A. F., Jimenez, R., Lahav, O., "Massive lossless data compression and multiple parameter estimation from galaxy spectra", *Monthly Notices of the Royal Astronomical Society*, 317, 965-972, 2000.

[10] Zhaoping, L., "Theoretical understanding of the early visual processes by data compression and data selection", *Network: Computation in Neural Systems*, 17, 301-334, 2006.

[11] Suarjaya, I.M.A.D., "A New Algorithm for Data Compression Optimization", *Int. J. Adv. Comp. Sci. and Appl.*, 3, 14-17, 2012.

[12] Adek, R.T., Ula, M., "A Survey on The Accuracy of Machine Learning Techniques for Intrusion and Anomaly Detection on Public Data Sets", *2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)*, 19-27, 2020.

[13] Thomas, R.N., Gupta, R., "A Survey on Machine Learning Approaches and Its Techniques" *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 1-6, 2020.

[14] Dixon, M.F., Halperin, I., Bilokon, P., "*Machine Learning in Finance from Theory to Practice*", Springer Nature Press, 2020.

[15] Bose, I., Mahapatra, R.K., "Business data mining a machine learning perspective.", *Inf. Management*, 39, 211-225, 2001.

[16] Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., Kim, M., Petit, J.E., Favez, O.,, "Meteorology-driven variability of air pollution (PM1) revealed with explainable machine learning", *Atmos. Chem. Phys.*, 21, 3919–3948, 2021.

[17] Vogt, N., "Machine learning in neuroscience", *Nature Methods*, 15, 33, 2018.

[18] Bektas, J., Bektas, Y., Kangal, E.E., "Integrating a novel SRCRN network for segmentation with representative batch-mode experiments for detecting melanoma", *Biomedical Signal Processing and Control*, 71, 103218, 2022.

[19] Ramakrishnan, R., von Lilienfeld, O.A., "*Chapter 5 in Reviews in Computational Chemistry*", Wiley-VCH, Weinheim, Germany, 225–256, 2017.

[20] Ibrahim, S.K., Ahmed, A., Zeidan, M.A.E., Ziedan, I.E., "Machine Learning Methods for Spacecraft Telemetry Mining", *IEEE Trans. Aerosp. Electron. Syst.*, 55, 1816-1827, 2019.

[21] Caldeira, J., Wu, W.L.K., Nord, B., Avestruz, C., Trivedi, S., Story, K.T., "DeepCMB: Lensing reconstruction of the cosmic microwave background with deep neural networks.", *Astron. Comput*. 28, 100307, 2019.

[22] Ntampaka, M. Trac, H., Sutherland, D. J., Battaglia, N., Póczos, B., Schneider, J., "A Machine Learning Approach for Dynamical Mass Measurements of Galaxy Clusters.", *Astrophys. J.*, 803, 50, 2015.

[23] Salti, M., Kangal, E.E., Aydogdu, O., "Evolution of CMB temperature in a Chaplygin gas model from deep learning perspective", *Astronomy and Computing*, 37, 100504, 2021.

[24] Salti, M., Kangal, E.E., "Deep learning of CMB radiation temperature", *Annals of Physics* 439, 168799, 2022.

[25] Kangal, E.E., Salti, M., Aydogdu, O., "Machine learning algorithm in a caloric view point of cosmology", *Phys. Dark Univ.*, 26, 100369, 2019.

[26] Escamilla-Rivera, C., Quintero, M.A.C., Capoziello, S., "A deep learning approach to cosmological dark energy models.", *JCAP*, 03, 008, 2020.

[27] Tilaver, H., Salti, M., Aydogdu, O., Kangal, E.E., "Deep learning approach to Hubble parameter.", *Comp. Phys. Commun.*, 261, 107809, 2021.

[28] Donalek, C., "*Supervised and unsupervised learning*", Astronomy Colloquia, California Institute of Technology, USA, 2011.

[29] Benvenuto, F., Piana, M., Campi, C., A. M. Massone, "A Hybrid Supervised/Unsupervised Machine Learning Approach to Solar Flare Prediction", *ApJ*, 853, 90, 2018.

[30] Arnott, D., Pervan, G., "Eight key issues for the decision support systems discipline", Decision Support Systems, 44, 657-672, 2008.

[31] Turban, E., Sharda, R., Delen, D., "*Decision Support and Business Intelligence Systems (9th edition)*", Pearson, London-UK, 2011.

[32] Hastie, T., Tibshirani, R., Friedman, J., "*The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd edition)*", Springer-Verlag, NY, USA, 2008.

[33] Kowsari, K., Meimandi, K.J., Heidarysafa, M., Sanjana, M., Laura, B., Brown, D., "Text Classification Algorithms: A Survey", *Information*, 10, 150, 2019.

[34] Harper, P.R., “A review and comparison of classification algorithms for medical decision making”, *Health Policy*, 71, 315-331, 2005.

[35] Kumar, R., Verma, R., “Classification Algorithms for Data Mining: A Survey”, *International Journal of Innovations in Engineering and Technology*, 1, 2319, 2012.

[36] Ilham, A., Khikmah, L., Ulumuddin, I., Indra I., “Long-term deposits prediction: a comparative framework of classification model for predict the success of bank telemarketing”, *IOP Conf. Series - Journal of Physics: Conf. Series*, 1175, 012035, 2019.

[37] Mustapha, S.M.F.D.S., Alsufyani, A., “Application of Artificial Neural Network and information gain in building case-based reasoning for telemarketing prediction”, *International Journal of Advanced Computer Science and Applications*, 10, 300-306, 2019.

[38] Tekouabou, S.C.K., Cherif, W., Silkan, H., “*A data modeling approach for classification problems: application to bank telemarketing prediction*”, Proceedings of the 2nd International Conference on Networking, Information Systems & Security (NISS19), Rabat-Morocco, 2019.

[39] Keles, A., Keles, A., “IBMMS Decision Support Tool for Management of Bank Telemarketing Campaigns”, *International Journal of Database Management Systems*, 17, 1, 2015.

[40] Kocoglu, F.O., Esnaf, S., “Machine Learning Approach and Model Performance Evaluation for Tele-Marketing Success Classification”, *International Journal of Business Analytics*, 9, 1-18, 2022.

[41] Halim, K.N.A., Jaya, A.S.M., Fadzil, A.F.A., “Data Pre-Processing Algorithm for Neural Network Binary Classification Model in Bank Tele-Marketing”, *International Journal of Innovative Technology and Exploring Engineering*, 9, 272-277, 2020.

[42] Shashidhara, B. M. Jain, S., Rao, V. D., Patil, N., Raghavendra, G.S., “*Evaluation of machine learning frameworks on bank marketing and Higgs datasets*”, Proceedings of Second International Conference on Advances in Computing and Communication Engineering, Dehradun-India, 2015.

[43] Wiemer, H., Drowatzky, L., Ihlenfeldt, S., “Data Mining Methodology for Engineering Applications (DMME)-A Holistic Extension to the CRISP-DM Model”, *Appl. Sci.*, 9, 2407, 2019.

[44] Jaggia, S., Kelly, A., Lertwachara, K., Chen, L., “Applying the CRISP-DM framework for teaching business analytics”, *Decision Sciences Journal of Innovative Education*, 18, 612-634, 2020.

[45] Mohamed, A., “Data Analysis for Telemarketing Case”, www.kaggle.com. Date of access: 04 Jan 2022.

[46] Alfeilat, H.A.A., Hassanat, A.B.A., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Salman, H.S.E., Prasath, V.B.S., “Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review”, *Big Data*, 7, 221-248, 2019.

[47] Deza, E., Deza, M.M., “*Encyclopedia of distances*”, Springer, 2009.

[48] Hall, M.A., “Correlation-based Feature Selection for Machine Learning”, PhD Thesis, The University of Waikato, Hamilton, NewZealand, 1999.

[49] Cunningham, S. J., Littin, J., Witten. I. H., “Applications of machine learning in information retrieval”, Technical Report 97/6, University of Waikato, 1997.

[50] Prasatha, V. B.S., Alfeilate, H.A.A., Hassanate, A.B.A., Lasassmehe, O., Tarawnehf, A.S., Alhasanatg, M.B., Salmane, H.S.E, “Effects of Distance Measure Choice on KNN Classifier Performance - A Review”, *e-Print:* 1708.04321v3, 2019.

[51] Cha, S.H., “Probabilistic, Statistical and Algorithmic Aspects of the Similarity of Texts and Application to Gospels Comparison”, *International Journal of Mathematical Models and Methods in Applied Sciences*, 1, 300-307, 2007.

[52] Jain, A., Zongker, D., “Feature selection: evaluation, application, and small sample performance”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 153-158, 1997.

[53] Pavlenko, T., “On feature selection, curse-of-dimensionality and error probability in discriminant analysis”, *Journal of Statistical Planning and Inference*, 115, 565-584, 2003.

[54] Myakalwar, A.K., Spegazzini, N., Zhang, C., Anubham, S.K., Dasari, R.R., Barman, I., Gundawar, M.K., “Less is more: Avoiding the LIBS dimensionality curse through judicious feature selection for explosive detection”, *Scientific Reports*, 5, 13169, 2015.

[55] Elkhani, N., Muniyandi, R.C., “Membrane computing inspired feature selection model for microarray cancer data”, *Intell. Data Anal.*, 21, 137-157, 2017.

[56] Last, M., Kandel, A., Maimon, O., “Information-theoretic algorithm for feature selection”, Pattern Recognition Letters, 22, 799-811, 2001.

[57] Muni, D.P., Pal, N.R., Das, J., “Genetic programming for simultaneous feature selection and classifier design”, *IEEE Trans. Syst. Man Cybern. Part B*, 36, 106-117, 2006.

[58] Strearns, S.D., “*On Selecting Features for Pattern Classifiers*”, 3rd International Conference on Pattern Recognition”, Colorado-CA, 1976.

[59] Pudil, P., Ferri, F.J., Novovicova, J., Kittler, J., “*Floating search methods for feature selection with nonmonotonic criterion functions*”, Proceedings of the 12th IAPR International Conference on Pattern Recognition, Vol. II-Conference B: Pattern Recognition and Neural Networks, Jerusalem-Israel, 1994.

[60] Pudil, P., Novovicova, J., Kittler, J., “Floating search methods in feature selection”, *Pattern. Recogn. Lett.*, 15, 1119-1125, 1994.

[61] Caruana, R., Niculescu-Mizil, A., “*Data mining in metric space: an empirical analysis of supervised learning performance criteria*”, Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining, Seattle-WA, USA, 2004.

[62] Li, H., Li, H., Wei, K., “Automatic fast double KNN classification algorithm based on ACC and hierarchical clustering for big data”, *Int. J. Commun. Syst.*, 31, e3488, 2018.

[63] Madray, I., Suire, J., Desforges, J., Madani, M.R., “Relative angle correction for distance estimation using K-nearest neighbors”, *IEEE Sensors Journal*, 20, 8155, 2020.

[64] Snchez-Crisostomo, J., Alejo, R., López-González, E., Valdovinos, R.M., Pacheco-Sánchez, J.H., "*Empirical analysis of assessments metrics for multi-class imbalance learning on the back-propagation context*" in “*Advances in Swerm Intelligence*”, Lecture Notes in Computer Science, Springer, 8795, 17-23, 2014.

[65] Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., “SMOTE: Synthetic Minority Over-sampling Technique”, *Journal of Artificial Intelligence Research*, 16, 321-357, 2002.

[66] Torgo, L., Ribeiro, R. P., Pfahringer, B., Branco, P., "*SMOTE for Regression*", Progress in Artificial Intelligence, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 8154, 2013.

[67] Fernández, A., Garcia, S., Herrera, F., Chawla, N.V., "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary", *Journal of Artificial Intelligence Research*, 61 863-905, 2018.

[68] Pavithra, P., Babu, S., *International Journal of Scientific Research and Engineering Development*, 2, 86-90, 2019.

[69] Jager, M., "*Improving data imbalance using Synthetic Minority Over-sampling (SMOTE)*", www.medium.com, Date of access: 04.06.2022.

[70] Probst, P., Bischl, B., Boulesteix, A.L., "Tunability: Importance of Hyperparameters of Machine Learning Algorithms", *e-Print:* 1802.09596, 2018.

[71] Wang, B, Gong, N. Z., "*Stealing Hyperparameters in Machine Learning*", 2018 IEEE Symposium on Security and Privacy (SP), San Francisco-CA, USA, 2018.

[72] Yang, L., Shami, A., "On hyperparameter optimization of machine learning algorithms: Theory and practice", *Neurocomputing*, 415, 295-316, 2020.

Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Research Article

ENHANCING MULTI-CLASS TEXT CLASSIFICATION WITH APRIORI-BASED FEATURE SELECTION

Maide Feyza ER*[1] ***Turgay Tugay BILGIN***[2]

[1] Bandirma Onyedi Eylul University, Faculty of Engineering and Natural Sciences, Department of Software Engineering, Balıkesir, Türkiye
[2]Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Bursa, Türkiye
[*] Corresponding author; mer@bandirma.edu.tr

Abstract: *In the field of Natural Language Processing, selecting the right features is crucial for reducing unnecessary model complexity, speeding up training, and improving the ability to generalize. However, the multi-class text classification problem makes it challenging for models to generalize well, which complicates feature selection. This paper investigates how feature selection impacts model performance for multi-class text classification, using a dataset of projects completed by TÜBİTAK TEYDEB between 2009 and 2022. The study employs LSTM, a deep learning method, to classify the projects into nine different industries based on various attributes. The paper proposes a new feature selection approach based on the Apriori algorithm, which reduces the number of attribute combinations considered and makes model training more efficient. Model performance is evaluated using metrics like accuracy, loss, validation scores, and test scores. The key findings are that feature selection significantly affects model performance, and different feature sets have varying impacts on performance.*

Keywords*: Feature Selection, LSTM, NLP, Text Classification*

Received: April 29, 2024 *Accepted: June 25, 2024*

1. Introduction

With the rapid development of the information age, advances in text mining and natural language processing have increased the capacity to extract meaning from large datasets. In this context, text classification has emerged as an important component of knowledge extraction. Multi-class text classification aims to classify text documents into multiple categories, and research in this area has focused on feature selection as a particularly important focus.

Feature selection is a critical factor affecting the performance of text classification models. Correct feature selection can lead to both computational savings and an improvement in the overall performance of the model [1]. This paper aims to delve deeper into the impact of feature selection on model performance in multiclass text classification. The studied effects have the potential to make the information extraction process more effective and improve efficiency in text mining applications.

Thirumoorthy et al. presented a feature selection technique based on the frequency distribution measure to address the high-dimensional feature space problem of text classifier accuracy reduction in the setting of a very high-dimensional feature space. In their study, the authors used SVM and Naive Bayes classifiers on two benchmark datasets to assess the efficacy of the suggested feature selection technique. They reported that, in terms of classification accuracy, the recommended feature selection method performed better than alternative feature selection strategies [2].

Amazal et al. on the other hand, proposed a distributed feature selection approach for large-scale multi-label textual big data using the weighted chi-square method implemented in the Hadoop framework. The suggested approach, which transforms multi-label data into single-label data, assigns weights to features based on the category term frequency and calculates the chi-square for each feature according to its weight. Experimental results have demonstrated that the proposed method is efficient, robust, and scalable when compared to state-of-the-art techniques [3].

Naik et al. highlighted the significance of text preprocessing, which entails organizing and cleaning data, in their work addressing the difficulties of processing and interpreting massive volumes of unstructured textual data. They emphasized how important it is to use Natural Language Processing (NLP) techniques to prepare data for analysis in order to extract useful information from text [4]. These techniques include language identification, tokenization, filtering, lemmatization, and stemming.

Dowlagar and Mamidi proposed a hybrid feature selection method that combines various filter-based feature selection techniques with the fastText classifier to obtain the necessary features for text classification. They observed a reduction in training time and a slight increase in accuracy in some datasets when feature selection methods were employed in conjunction with neural networks [5].

Hussain and colleagues proposed a novel feature ranking score called the Differential Mutual Information (DMI) score and an avant-garde technique called Non-Redundant Feature Selection (NRFS) in their work to solve the shortcomings of conventional feature selection methods in text data. In comparison to previous state-of-the-art methods, the suggested method was shown to produce superior micro-F1 classification scores and to exhibit more resilience against label noise, especially in situations when there were few picked features [6].

Belkarkor and colleagues's research centered on the difficulty of handling large amounts of data in text processing and the significance of feature selection in machine learning algorithms for text categorization. With the use of three reference document collections and the NB classifier, the study assessed the effectiveness of Genetic Algorithm in feature selection by contrasting it with alternative filtering techniques. According to the study, Genetic Algorithm fared better in text categorization than other filtering techniques in terms of efficiency and accuracy [7]. It was noted that the best feature selection strategy differed for every dataset in Zheng's paper, which compared various feature selection techniques. Nonetheless, some techniques regularly yielded valuable data for class categorization, and chi-square was widely acknowledged as the best technique [8].

In order to improve the efficacy of feature selection in text mining and machine learning, Tang et al. presented a feature selection technique that uses two to five-way interactions to account for high-order feature interactions. By breaking down the mutual information-based feature selection problem into lower-order interactions, the strategy depends on an effective measurement for predicting interaction terms. They claimed that by taking into account high-level feature interactions, their suggested approach performed better when it came to feature selection for text categorization tests [9].

By analyzing various feature selection techniques and comprehending the variances in the performance of multi-class text classification models, this study aims to give academics important insights into overcoming difficulties presented by multi-featured datasets. This paper will examine pertinent research in natural language processing and text mining, highlighting the critical role feature selection plays from a strategic standpoint. Next, the emphasis will be on the findings of research investigations carried out to provide insight into how feature selection affects text classification model performance.

In this context, an overview of the data set is presented in the Materials and Methods section of the article. Following, text preprocessing, feature selection algorithm and model training are explained respectively. The findings are discussed in the Results and Discussion section and finally the Conclusion is given in the last section.

2. Materials and Methods

This section provides a detailed explanation of the working steps. It consists of four main headings in total: Data Set Overview, Text Preprocessing, Feature Selection Algorithm, and Model Training. The method proposed in the study is presented in Figure 1.

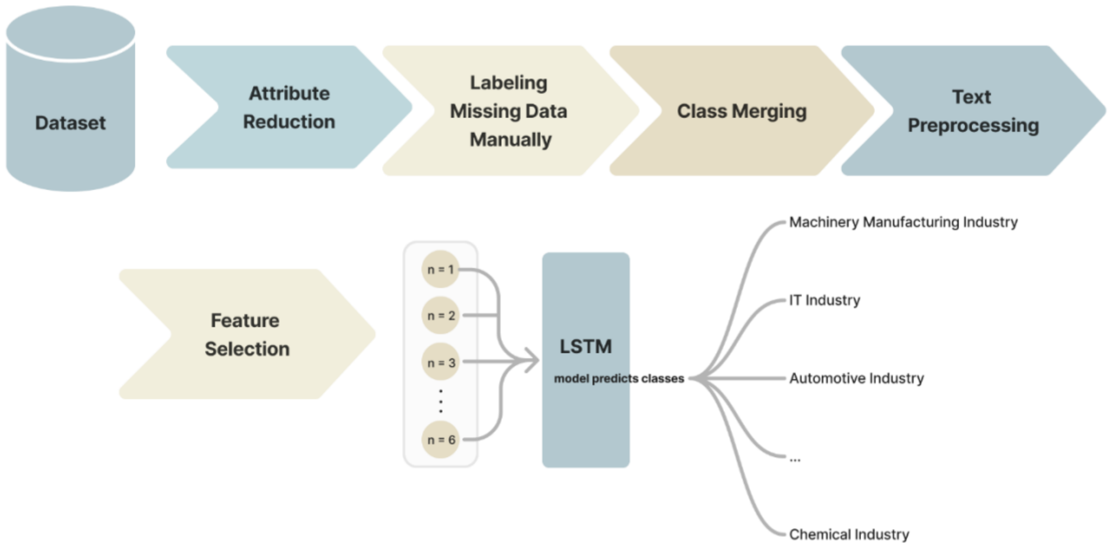

Figure 1. The proposed method in the study.

2.1. Dataset Overview

Initially, we collected the data via an RPA Tool from the TEYDEB Project Evaluation and Monitoring System portal of the TUBITAK website. The compilation comprises information about TUBITAK TEYDEB projects completed in the years 2009 through 2022. Researchers may access the details of completed projects on the TUBITAK TEYDEB website [10]. There are 1672 records with 11 attributes in the collection. Table 1 enumerates each attribute in the dataset and the number of items that correspond to that attribute.

Table 1. The dataset's attributes and the quantity of entries for each attribute.

Attributes	Number of Entries
Project Name	1672
Organization Name	1672
Organization Province	1672
Project Year	1672
Keywords	1406
Project Start-End Date	1672
Scientific Technological Activity Area	1656
Industry that R&D Activities are Conducted	1127
Industry that Project Outputs are Utilized	1038
Project Summary	671
Project Objectives	1194
Technical Specifications of Project Outputs	1449

It is evident from Table 1 that certain features in the dataset will not help to improve the classification success rate. The “Organization Name”, “Organization Province”, “Project Year”, and “Project Start-End Date” are examples of these features. They have been removed from the dataset.

Table 1 also shows that certain attributes in the dataset have missing data. There are two main methods for handling missing data. The first and easiest option is to remove missing data; the second method uses random, statistical, or k-nearest neighbor algorithms to fill in the missing data. However, it should be noted that these methods may have adverse effects on the dataset and are not universally applicable to every situation. Deleting missing data can lead to further reduction in small datasets, causing data loss. Additionally, filling in missing data is generally applicable only to numerical values.

The "Industry that Project Outputs will be used" attribute will serve as the basis for classification, hence any missing data in this attribute has to be manually labeled because removing it would result in a significant loss of data. The “2.3 Feature Selection” section provides an explanation of the method used to select the solution that is deemed most appropriate for the dataset's characteristics. Manually filling in the missing values for the other attributes is not feasible. After manually filling, the number of classes for the attribute's classification and the total number of data in these classes were analyzed. Table 2 provides the number of projects grouped by the Industries.

Table 2. Number of projects grouped by the industries

The industry that Project Outputs will be Utilized	Number of Data
Machinery Manufacturing Industry	354
IT Industry	217
Automotive Industry	175
Electrical & Electronics Industry	130
Biomedical Industry	112
Energy Industry	103
Food Industry	83
Material Industry	83
Textile Industry	69
Chemical Industry	66
Defense Industry	43
White Goods Industry	37
Metallurgical Industry	35
Pharmaceutical Industry	30
Telecommunications Industry	30
Agriculture Industry	29
Environmental Technologies Industry	21
Ship and Maritime Industry	13
Aerospace Industry	11
Mining Industry	10
Livestock Industry	10
Aquaculture Industry	9
Ceramic/Earth Products	2

There are a total of 23 classes in the dataset, as Table 2 shows, and the distribution of the number of entities is not balanced. There are many classes available for classification despite the small size of the dataset. This is an important issue that can make the classifier less effective. The merging of closely related classes has been considered as an improvement technique, even if it would not be possible to fully solve this issue. As a result, there would be fewer classes and more entities in each class.

Data from the Agriculture and Livestock Industry are combined with the Food Industry. In addition, the Pharmaceuticals-Pharmaceuticals Industry was merged with the Chemicals Industry. The Metallurgy Industry was merged with the Materials Industry. The Telecommunications Industry was merged with the Information Technology Industry. Lastly, the Household Appliances Industry with the Electrical and Electronics Industry merged. Classes with less than 90 entities have been removed from the dataset after these merges. As a result, the final situation regarding industries and number of projects is shown in Table 3.

Table 3. The dataset's attributes and the quantity of entries for each attribute.

Industries that Project Outputs will be Utilized	Number of Data
Machinery Manufacturing Industry	354
IT Industry	247
Automotive Industry	175
Electrical & Electronics Industry	167
Food Industry	122
Material Industry	118
Biomedical Industry	112
Energy Industry	103
Chemical Industry	96

After merging, classes and their respective data densities are displayed in Figure 2. The distribution of the classes is found to be more uniform than it was in the original data set. It is clear that the classifiers will perform better in this way.

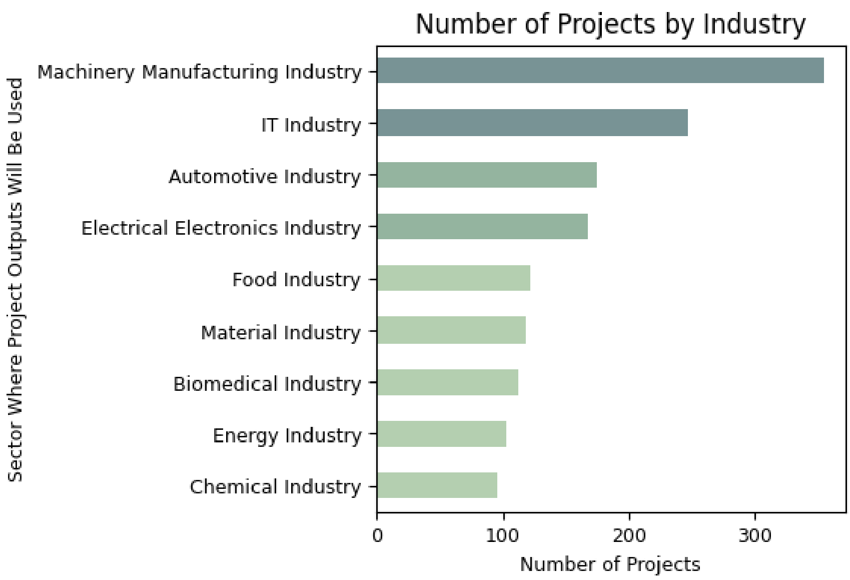

Figure 2. Barchart shows class labels and the data density of classes after merging.

2.2. Text Preprocessing

Text preprocessing, also known as the process of preparing raw text for natural language processing, aims to extract valuable information from the text by removing attributes that are irrelevant to the analysis. In this context, text preprocessing is critical in optimizing text for analysis by removing non-essential attributes from the dataset.

Before the dataset can be classified, we should remove user-generated textual noise and overcome difficulties specific to the Turkish language. Since all texts in the dataset will be converted into numeric values before classification, making all texts with the same meaning identical is the main goal of text preprocessing. In this context, in the preprocessing stage, Python's 're' module was utilized to check for certain patterns in the text and replace them with the desired patterns. This module enables checking whether a string of characters matches a specific regular expression. A regular expression specifies a pattern that matches a set of characters in itself [11].

Regular Expressions in Python, provided through the re-module, constitute a specialized parser. This parser allows for the establishment of rules to match specific potential word sequences. These rules can be utilized to check whether a string of characters matches a desired pattern or if there is a match for a specific pattern. Additionally, regular expressions can be used to modify a string or split it in various ways [11].

The first stage of text preprocessing involves converting all letters to lowercase. The conventional method of converting to lowercase proves inadequate in transforming Turkish characters not found in the Latin alphabet. For instance, the letter "I" is converted to lowercase as "i". To overcome this issue, specific regex rules have been defined for all Turkish-specific characters. After converting to lowercase, the text preprocessing process is completed by removing all characters and numbers except alphanumeric characters.

2.3. Feature Selection Algorithm

After removing redundant attributes from the dataset and manually filling in missing data in the "Industry that Project Outputs will be used" attribute, the final version of the dataset is provided in Table 4.

Table 4. Total number of entries for each attribute after removing redundant attributes.

Attributes	Number of Entries
Project name	1672
Keywords	1406
Scientific Technological Field of Activity	1656
Industry that R&D Studies will be carried out	1127
Industry that Project Outputs will be used	1672
Summary of the Project	671
The goal of the project	1194
Technical Specifications of Project Outputs	1449

As seen from Table 4, missing data still exists in all attributes except "Project Name" and "Industry that Project Outputs will be used". A classification model can be trained with the single attribute: "Project Name". However, using a single attribute will significantly decrease the model performance. Therefore, in this study, all attributes are used in various combinations while training, to investigate which attribute has a positive impact on the classifier performance. All of the attribute combinations include the "Project Name" attribute. In this context, all subsets of the attribute combinations containing “Project Name” were employed to train the model. With this approach, we intended to reduce the effect of missing data problems on model performance and to observe the effect of various attribute combinations on performance.

In the first attempt of model training, the “Project Name” attribute will be used. In the second attempt, all two element combinations containing the “Project Name” attribute will be used to train the model. In the next attempt, all three-element combinations containing the pair that provides the highest model performance among the two-element subsets will be used for model training. The same cycle will continue until there is only one subset containing all attributes.

Our proposed method for the feature selection is given in Figure 3. The approach is similar to the Apriori algorithm used for "association analysis". When the number of features is n=1, the model will be trained only with the “Project Name” feature. When n=2, all two-element combinations containing the “Project Name” attribute will be used. The subset with the highest performance from the two-element combinations will be selected, and all three-element subsets containing these two selected features will be used. In this way, all combinations including all elements will be subjected to model training.

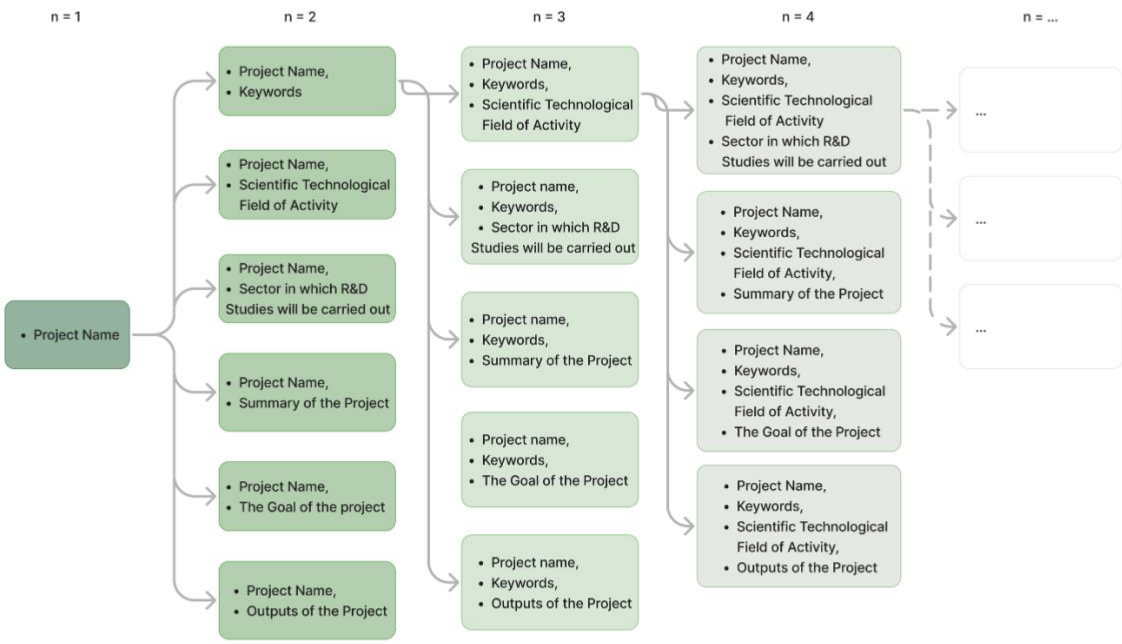

Figure 3. A representative illustration of feature selection based on the Apriori Algorithm. For illustrative purposes figure is drawn for “Project Name” and “Keywords” attributes.

To adapt the Apriori algorithm for feature selection, our proposed algorithm steps are given below:

1. **Initialization**: Start with individual features as item sets.
2. **Iteration**:
   - Generate candidate feature sets by combining existing features.
   - Evaluate the classification model accuracy using each candidate feature sets.
   - Prune candidate feature sets that do not improve classification accuracy above a predefined threshold.
3. **Termination**: Repeat the iteration until no more candidate feature sets can be generated or until the desired number of features is reached.

The key difference from traditional Apriori is in the evaluation step, where instead of measuring the frequency of item sets, we measure the classification accuracy of feature sets. The pruning step remains similar, where feature sets that do not meet the predefined accuracy threshold are discarded.

Considering only feature sets that meet the predefined accuracy threshold will significantly reduce the number of calculations required to generate new subsets. The total number of subsets for a set of N elements can be calculated using the formula for the power set. Let N be the number of elements in the set. The total number of subsets S(N) for a set of N elements is given by Eq. (1).

$$S(N) = 2^N \tag{1}$$

Since the total number of features in the dataset is 7, the total number of subsets is 128. This implies that 128 combinations of feature sets should be used for model training. On the other hand, in our proposed Apriori-like method, new subsets are generated only from the top-performing subsets. Thus, the number of trials for model training is significantly reduced.

According to the suggested feature selection approach, there are 5 three-element subsets when there are 6 two-element subsets that contain the “Project Name”. The total number of viable combinations reduces by 1 as the subset's element count rises. In this scenario, it is possible to compute the total number of subgroups using Equation (2).

$$\frac{N \times (N+1)}{2} \tag{2}$$

There are 21 subsets with “Project Name” with more than one element. Only the subset with the attribute “Project Name” will be used for model training among the subsets containing a single element. As a result, a total of 22 distinct attribute sets will be used to train the model. This reduces the number of model trials from 128 to 22, yielding faster and more efficient results.

2.4. Model Training

In this study, Long Short-Term Memory (LSTM) is used for model training. LSTM was introduced as a solution to the vanishing or exploding gradient problems that arise in Recurrent Neural Networks (RNN), particularly when learning long-term dependencies [12]. Therefore, LSTMs represent a specialized type of RNN capable of learning long-term dependencies and retaining information over extended periods by default [13].

LSTM utilizes short-term and long-term memory cells to determine the importance of data. If the data is deemed important, it is stored in memory and fed back into the network; otherwise, the data is forgotten to clear the memory. LSTM has a chain structure consisting of identical cells that repeat each other [13]. The chain structure of LSTM, consisting of a series of recurring connected cells, is given in Figure 4. Each cell is designed in a special way to store and transmit information.

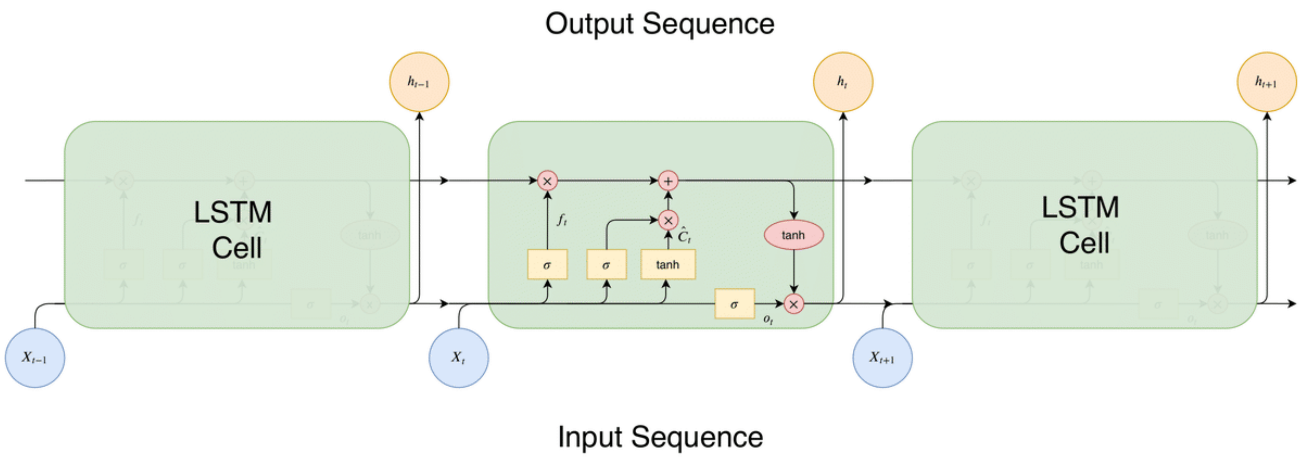

Figure 4. Chain structure of LSTM network.

The internal architecture of an LSTM cell is given in Figure 5. The basic components of an LSTM cell consist of the Cell State and three main gates: Input Gate, Output Gate and Forget Gate. While the Input Gate i_t decides when to read the data, the Forget Gate f_t is used to reset the contents of the cell. On the other hand, Output Gate o_t transfers the output of the cell to the next cell. Lastly, the Cell State, which is the main component of the data flow, ensures that the data flows forward without changing. This behavior of the Cell State reduces the vanishing gradient problem and provides a better understanding of long-time dependencies [13].

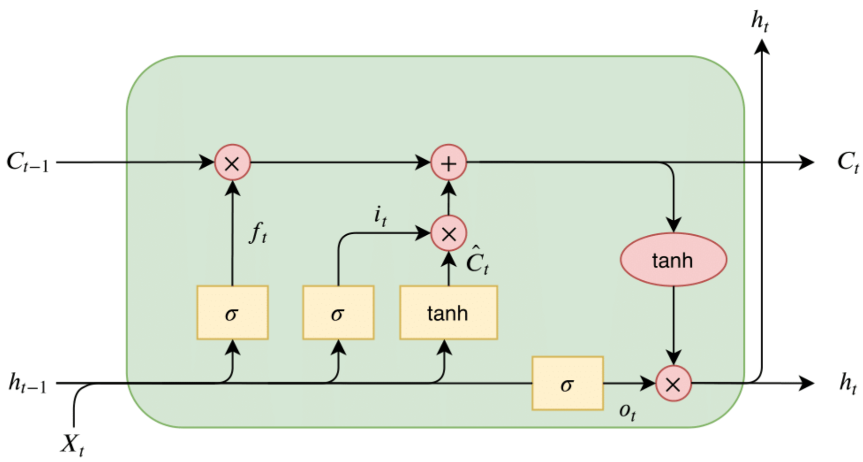

Figure 5. Cell architecture of LSTM.

As seen in Figure 5, a cell receives three inputs: X_t, (h_{t-1}) and (C_{t-1}). X_t is the input that feeds the cell at the current time t, (h_{t-1}), which is called the hidden state, is the output of the previous cell and (C_{t-1}) is the cell state of the previous cell. Subsequently, it modifies the previous memory cell (C_{t-1}) to produce two new outputs, (h_t) and (C_t). C_t is the newly updated memory, known as the cell state, while h_t is the output of the current LSTM cell, known as the hidden state. At each time step, these two states, (C_t) and (h_t) are transferred from the current cell to the next cell. Likewise, the previous cell state (C_{t-1}) and the previous hidden state (h_{t-1}) were transferred from the previous cell to the current cell at time step *t-1* [13].

The LSTM cell receives inputs, X_t and (h_{t-1}), and feeds them to three LSTM gates. These inputs pass through the sigmoid, which is the activation function of all three gates, and produce numbers between "0" and "1" as output. Therefore, a value of 0 ('off') will not allow anything to pass through the gate, while a value of 1 ('on') will allow everything to pass through the gate. Based on this, the equation of the input gate is given in equation (3), the forget/hidden gate is given in equation (4), and the equation of the exit gate is given in equation (5).

$$i_t = \sigma(X_t V_{xi} + h_{t-1}W_{hi} + C_{t-1}W_{ci} + b_i) \quad (3)$$
$$f_t = \sigma(X_t V_{xf} + h_{t-1}W_{hf} + C_{t-1}W_{cf} + b_f) \quad (4)$$
$$o_t = \sigma(X_t V_{xo} + h_{t-1}W_{ho} + C_{t-1}W_{co} + b_o) \quad (5)$$

V_x, W_h and W_c in the equations are the weight vector associated with X_t, (h_{t-1}) and (C_{t-1}), respectively, while b is the bias weight vector [12]. According to Equation (3), the i_t vector determines how much of the input needs to be updated. The term $X_t V_{xi}$ represents the contribution of the current input to the LSTM cell. On the other hand, the term $h_{t-1}W_{hi}$ represents the contribution of the hidden state to the cell at the previous time. Lastly, the term $C_{t-1}W_{ci}$ represents the contribution of the cell state to the cell at the previous time. The bias term b_i is an additional learnable variable and makes the model more flexible and the sigmoid function σ, is used to determine how much of each entry to update.

As seen in Equation (4), the vector f_t determines how much of the cell state is forgotten. The sigmoid function compresses the output of this gate to a value between 0 and 1, which calculates how much of the cell state will be forgotten. Finally, in Equation (5), o_t determines how much of the cell state is used as output.

In order to calculate the cell state, C_t, the candidate cell state, $\hat{C}_t$ must first be calculated. The candidate cell state is calculated by combining the current input and hidden state vectors. The candidate cell state refers to the potentially updated state of the cell, but not yet filtered by forget and output gates. The candidate cell state and new cell state formula is calculated as follows:

$$\hat{C}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (6)$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \quad (7)$$

In Equation 6, $\hat{C}_t$ is the candidate cell state, and this vector is calculated by combining the input vector x_t and the previous hidden state vector h_{t-1}. W_{xc} and W_{hc} in Equation 6 are the weight matrices, while b_c is the bias term. The hyperbolic tangent function $\tanh$ limits the candidate cell state and scales the output between -1 and 1.

Once the candidate cell state is calculated, the new cell state can now be calculated. The new cell state is updated by the combination of the previous cell state and the candidate cell state, as given in Equation 7. The f_t in the formula is the output of the forget gate and multiplied by C_{t-1} determines how much of the previous cell state will be forgotten. The i_t is the output of the input gate and is multiplied by $\hat{C}_t$ to determine how much of the candidate cell state will be added to the cell state. The $\odot$ sign in Equation 7 represents element-wise multiplication.

Finally, the hidden state, h_t is calculated. The formula for the hidden state calculated using the output gate and the updated cell state is given in Equation 8. Accordingly, the cell state is passed through the tanh function and subjected to element-wise multiplication with the output gate. Thus, the hidden state is compressed into the (-1, 1) range.

$$h_t = o_t \odot \tanh(C_t) \quad (8)$$

LSTM is well suited for time series forecasting as well as other problems that require temporal memory, as it can learn long-term dependencies without the problem of exploding/vanishing gradients [12]. The key reason for the success of LSTMs in natural language processing lies in their ability to handle long-term dependencies effectively. Sentences in natural language often involve long-term dependencies, where the meaning of a word is strongly related to previous and subsequent words. LSTMs can manage long-term dependencies more effectively through mechanisms such as the forget gate, input gate, and output gate.

3. Results and Discussion

In this study, LSTM, one of the deep learning methods, is used to predict the value of the “Industry that Project Outputs will be used” field. In order to find out which attribute performs better, an Apriori based method was used to train models with different attribute combinations. Model training with only one attribute was performed only for the “Project Name” attribute. The metrics of the model are given in Table 5.

The performance of the model has been evaluated using the metrics Loss, Accuracy, Validation Loss, Validation Accuracy, Test Loss, and Test Accuracy. Accordingly, Accuracy is the ratio of the correctly predicted examples to the total predictions, while Loss measures how erroneous the model's predictions are. Validation Loss indicates the errors made by the model on the validation dataset, whereas Validation Accuracy is the ratio of correctly predicted examples in the validation dataset. Test Loss measures the errors made by the model on the previously unseen test dataset, and Test Accuracy is the ratio of correctly predicted examples in the test dataset [14].

An accuracy of 0.8900 and an error value of 0.4322 were attained, as Table 5 demonstrates. Simply looking at these findings could give the impression that the model operates fairly effectively. However, a notable rise in error and fall in accuracy is seen when looking at the accuracy and error metrics for the Validation and Test sets. This indicates that the model has memorized the data, suggesting overfitting. Indeed, it highlights that the “Project Name” attribute alone has a very weak generalization ability for the model. Therefore, in the next step, other attributes will be used for training alongside the “Project Name” attribute.

Table 5. Metrics of the model trained with the “Project Name” attribute.

Metrics	Values
Loss	0.4322
Accuracy	0.8900
Validation Loss	2.1488
Validation Accuracy	0.2889
Test Loss	2.3862
Test Accuracy	0.2600

Table 6 displays the model's results for each two-element attribute subset that contains “Project Name”. As shown in the table, the subset titled “Project Name – Project Summary” obtained the best accuracy and lowest loss rate. Nevertheless, it is noted that in terms of test loss, test accuracy, validation accuracy, and validation accuracy numbers, this combination does not produce the best results. Poor performance on the validation and test set compared to the training set is indicative of overfitting. When these metrics are examined on the test set, it becomes clear that this attribute pair does not effectively satisfy the generalization capability of the model.

Table 6. Metrics of all two-element attribute subsets containing "Project Name".

Combination of Attributes Used	Loss	Accuracy	Validation Loss	Validation Accuracy	Test Loss	Test Accuracy
Project Name, Keywords	0.6935	0.7866	1.8048	0.3630	1.9846	0.3333
Project Name, Scientific and Technological Field of Activity	0.6773	0.8222	1.7980	0.4593	1.7488	0.5200
Project Name, Industry that R&D Studies will be carried out	0.4278	0.8784	2.1168	0.4296	1.9482	0.4600
Project Name, Project Summary	0.4197	0.9007	2.5104	0.3259	2.0752	0.3800
Project Name, Purpose of the Project	0.9899	0.6857	2.2289	0.2963	1.9136	0.2600
Project Name, Outputs of the Project	0.6692	0.7974	2.7115	0.1407	2.3502	0.2733

The "Project Name - Scientific Technological Activity Area" feature combination performed well in both training and test/validation sets, producing more balanced results. With an accuracy value of 0.8222, this combination not only performed quite well, but also produced the highest validation and testing performance among all two-element feature subsets. Therefore, in the next step, when creating three-element subsets, all subsets must contain the pair "Project Name - Scientific Technological Activity Area".

The model metrics for all three-element attribute subsets containing the pair "Project Name – Scientific Technological Activity Area" are given in Table 7. Here, it is observed that all accuracy values in Table 7 yield quite good results compared to the ones in Table 6. However, examining the validation and test values reveals declines similar to those in Table 6.

Table 7. Model metrics for All Three-Element Attribute Subsets Containing the Pair "Project Name – Scientific Technological Activity Area."

Combination of Attributes Used	Loss	Accuracy	Validation Loss	Validation Accuracy	Test Loss	Test Accuracy
Project Name, Scientific and Technological Field of Activity, Keywords	0.4130	0.8859	2.2889	0.3407	2.4073	0.3267
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out	0.4828	0.8412	1.7245	0.5259	1.5907	0.5000
Project Name, Scientific and Technological Field of Activity, Summary of the Project	0.5074	0.8395	1.9842	0.3778	2.1492	0.3667

Table 7. Continued

Combination of Attributes Used	Loss	Accuracy	Validation Loss	Validation Accuracy	Test Loss	Test Accuracy
Project Name, Scientific and Technological Field of Activity, Purpose of the Project	0.7008	0.8147	2.0090	0.3333	1.8647	0.3933
Project Name, Scientific and Technological Field of Activity, Outputs of the Project	0.6026	0.8420	2.6269	0.2741	2.8459	0.2933

The trio “Project Name, Scientific Technological Activity Area, Keywords” has the lowest error of 0.4130 and the best accuracy of 0.8859 for three-element attribute subsets. Nonetheless, the attribute trio “Project Name, Scientific Technological Activity Area, Industry that R&D Activities are Conducted” yields the best results when it comes to validation and test performance. The power of the model for generalization can be seen more clearly by this trio. Thus, this trio will be employed to generate four-element attribute subsets.

The attribute trio “Project Name, Scientific Technological Activity Area, Industry that R&D Activities are Conducted,” when compared to the pair “Project Name – Scientific Technological Activity Area,” yielded better performance in all metrics except for test accuracy. As the number of attributes increases, the accuracy of the model increases and the error rate decreases. This suggests that the model may be learning more effectively with additional informative attributes, fitting better to the dataset.

The four-element attribute subsets containing the attribute trio “Project Name, Scientific Technological Activity Area, Industry that R&D Activities are Conducted” are provided in Table 8. The highest accuracy value is found in the subset that includes the “Project Purpose” attribute, which surprisingly has the lowest validation accuracy. This suggests that the model may be experiencing an overfitting problem, despite achieving the highest accuracy in this particular subset.

Table 8. Model metrcis for All Four-Element Attribute Subsets Containing the Trio “Project Name – Scientific Technological Activity Area – Industry that R&D Activities are Conducted”.

Combination of Attributes Used	Loss	Accuracy	Validation Loss	Validation Accuracy	Test Loss	Test Accuracy
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out, Keywords	0.4146	0.8958	1.9712	0.3556	2.1429	0.3667
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out, Summary of the Project	0.6793	0.8056	2.1528	0.2519	2.1550	0.1800
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out, Purpose of the Project	0.3823	0.9165	2.5463	0.1852	2.4456	0.2533
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out, Outputs of the Project	0.6752	0.7891	2.3601	0.2519	2.1120	0.2800

Table 9. Model metrics for All Five-Element Attribute Subsets Containing the Quartet "Project Name – Scientific Technological Activity Area – Industry that R&D Activities are Conducted – Keywords".

Combination of Attributes Used	Loss	Accuracy	Validation Loss	Validation Accuracy	Test Loss	Test Accuracy
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out, Keywords, Summary of the Project	0.6018	0.8131	1.8527	0.3630	2.0945	0.3600
Project Name, Field of Scientific Technological Activity, Industry that R&D Studies will be carried out, Keywords, Purpose of the Project	0.9220	0.7328	2.3042	0.2963	2.2629	0.2933
Project Name, Field of Scientific and Technological Activity, Industry that R&D Studies will be carried out, Keywords, Outputs of the Project	0.8908	0.7411	2.5708	0.3630	2.4176	0.3533

In general, the performances reflected in Table 8 appear to be significantly lower compared to previous results. The relatively best performance —both in terms of accuracy and validation and test sets— is observed in the subset containing the "Keywords" attribute. Therefore, subsets with five elements will be generated using this attribute set.

In Table 9, model metrics are provided for all five-element attribute subsets containing the quartet "Project Name – Scientific Technological Activity Area – Industry that R&D Activities are Conducted – Keywords". The subset containing the "Project Summary" attribute demonstrates relatively superior performance in terms of accuracy and error values, as well as in test and validation metrics. Therefore, the six-element subsets are constructed based on this quintet subset. The corresponding six-element subsets are also given in Table 10.

Table 10. Model metrics for All Six-Element Attribute Subsets Containing the Quintet "Project Name – Scientific Technological Activity Area – Industry that R&D Activities are Conducted – Keywords – Project Summary".

Combination of Attributes Used	Loss	Accuracy	Validation Loss	Validation Accuracy	Test Loss	Test Accuracy
Project Name, Field of Scientific Technological Activity, Industry that R&D Studies will be carried out, Keywords, Summary of the Project, Purpose of the Project	1.0779	0.6311	2.3087	0.1778	1.7953	0.3667
Project Name, Field of Scientific Technological Activity, Industry that R&D Studies will be carried out, Keywords, Summary of the Project, Outputs of the Project	0.1433	0.9760	2.7160	0.2593	2.5852	0.3133

Table 10 shows that, with an accuracy of 0.9760, the subset holding the “Project Outputs” attribute achieves very high model performance. The model is overfitting, nevertheless, as seen by its comparatively poor validation accuracy of 0.2593 and test accuracy of 0.3133.

At last, Table 11 presents the model that was trained utilizing each attribute in the dataset. As seen in Table 11, there is an increase in model accuracy. However, the lack of improvement in the test and validation metrics indicates that the memorization problem encountered in previous models exists here as well.

Table 11. Table contains the metrics of the model trained using all attributes in the dataset.

Metrics	Values
Loss	0.3477
Accuracy	0.9330
Validation Loss	2.6620
Validation Accuracy	0.2370
Test Loss	2.9152
Test Accuracy	0.2600

4. Conclusion

In this study, we delved into the critical aspect of feature selection in the context of multi-class text classification. Our study centered around a comprehensive dataset comprising TUBITAK TEYDEB projects, and we sought to unravel the impact of different attribute subsets on model performance. To tackle this challenge, we employed Long Short-Term Memory (LSTM), a deep learning architecture known for its ability to handle sequential data and capture long-term dependencies. Our goal was to classify the projects into nine distinct output industries based on relevant features.

The crux of our approach lay in feature selection. We proposed a novel method grounded in the Apriori algorithm, which systematically pruned the attribute space. By doing so, we aimed to strike a balance between model efficiency and predictive accuracy.

The dataset used in the study contains information about TUBITAK TEYDEB projects. The dataset can be enriched with information obtained from other web sources. This could potentially lead to improved performance. The proposed feature selection method aims to reduce computation costs by selecting new feature sets from the subsets that yields the best performance. However, this does not guarantee finding an optimal feature set. Further research will focus on developing more advanced and effective feature selection methods.

Our findings underscored the pivotal role of feature selection. Among various attribute combinations, the trio comprising “Project Name”, “Scientific Technological Activity Area”, and “Industry that R&D Activities are Conducted” emerged as the most potent. This subset consistently yielded superior model performance across multiple evaluation metrics, including accuracy, loss, validation accuracy, and test accuracy. Our study highlights that judicious feature selection significantly enhances the accuracy and generalization ability of multi-class text classification models. Researchers and practitioners alike should consider these insights when designing robust and efficient systems for real-world applications.

The LSTM model for the multiclass text classification problem is the main topic of this research. There are more sophisticated and recent models, but LSTM is a model that can manage sequential data and long-term dependencies well. It can be utilized in conjunction with these new models in further studies.

Ethical Statement

Our study does not cause any harm to the environment and does not involve the use of animal or human subjects. Therefore, it was not necessary to obtain an Ethics Committee Report.

Conflict of Interest

The authors declare no conflicts of interest.

Authors' Contributions

M.F.E: Investigation, Methodology, Preprocessing, Writing (%60)
T.T.B: Conceptualization, Data collection, Methodology, Formal analysis (%40).
All authors read and approved the final manuscript.

References

[1] Dogra, V., Singh, A., Verma, S., Kavita, Jhanjhi, N. Z., Talib, M. N., Understanding of data preprocessing for dimensionality reduction using feature selection techniques in text classification, in: *Intelligent Computing and Innovation on Data Science: Proceedings of ICTIDS*, Springer, Singapore, pp. 455-464, 2021.

[2] Thirumoorthy, K., Muneeswaran, K., "Feature selection for text classification using machine learning approaches." *National Academy Science Letters*, 45(1), 51-56, 2022.

[3] Amazal, H., Ramdani, M., Kissi, M. (2020). "Towards a feature selection for multi-label text classification in big data." *Proceedings of Smart Applications and Data Analysis: Third International Conference*, Marrakesh, Morocco, June 25–26, 2020, pp. 187-199

[4] Naik, D. A., Mythreyan, S., Seema, S., "Relevance Feature Discovery in Text Mining Using NLP". in: *3rd International Conference for Emerging Technology*, IEEE, pp. 1-6, 2022.

[5] Dowlagar, S., Mamidi, R. "Does a Hybrid Neural Network Based Feature Selection Model Improve Text Classification?", *arXiv preprint arXiv:*2101.09009, 2021. https://doi.org/10.48550/arXiv.2101.09009

[6] Hussain, S. F., Babar, H. Z. U. D., Khalil, A., Jillani, R. M., Hanif, M., Khurshid, K. "A fast non-redundant feature selection technique for text data", *IEEE Access*, 8, 181763-181781, 2020.

[7] Belkarkor, S., Hafidi, I., Nachaoui, M., Feature Selection for Text Classification Using Genetic Algorithm, in: *International Conference of Machine Learning and Computer Science Applications*, Cham: Springer Nature Switzerland, pp. 69-80, 2022.

[8] Zheng, W., "A comparative study of feature selection methods." *International Journal on Natural Language Computing*, 7(5), 01-09, 2018.

[9] Xiaochuan, T., Xiaochuan, T., Yuanshun, D., Yanping, X., "Feature selection based on feature interactions with application to text categorization." *Expert Systems with Applications*, 120, 207-216, 2019. doi: 10.1016/J.ESWA.2018.11.018

[10] TÜBİTAK TEYDEP Proje Değerlendirme ve İzleme Sistemi (2023, Nov. 23). Tamamlanmış Proje Sorgula [Online]. Available: https://eteydeb.tubitak.gov.tr/

[11] re — Regular expression operations. (2024, Jan. 05). Regular expression operations [Online]. Available: https://docs.python.org/3/library/re.html

[12] Van Houdt, G., Mosquera, C., Nápoles, G., "A review on the long short-term memory model", *Artificial Intelligence Review*, 53(8), 5929–5955, 2020. https://doi.org/10.1007/s10462-020-09838-1

[13] PO, U., Udanor, C. N., Bakpo, F. S., "Deep Learning Algorithms for predicting the geographical locations of Pandemic Disease Patients from Global Positioning System (GPS) Trajectory Datasets". https://doi.org/10.21203/rs.3.rs-2770308/v1

[14] Rainio, O., Teuho, J., Klén, R. "Evaluation metrics and statistical tests for machine learning". *Scientific Reports*, 14(1), 6086. 2024.

Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Research Article

ASSESSING THE POTENTIAL PUBLIC HEALTH RISK OF *Clostridium botulinum* TOXIN GENES IN CANNED FOOD: A LABORATORY EXPERIENCE IN TÜRKİYE

Hisamettin DURMAZ*[1] *Akın YİGİN*[2] *Serap Kılıç ALTUN*[1] *Mehmet Emin AYDEMİR*[1]

[1]Dep. of Food Hygiene and Technology, Faculty of Veterinary Medicine, Harran University, Sanlıurfa, Türkiye
[2]Department of Genetics, Faculty of Veterinary Medicine, Harran University, Sanlıurfa, Türkiye.
*Corresponding author: aydemiremin23@harran.edu.tr

***Abstract**: Due to the ability to synthesize a highly potent neurotoxin, Clostridium botulinum (C. botulinum) poses an important risk to food safety. Foodborne botulism is a neuroparalytic disease caused by the ingestion of neurotoxins produced by C. botulinum. Botulism commonly occurs as a result of consuming canned foods. In this study, the aim was to investigate the presence of C. botulinum and detect the genes producing A, B, E, and F toxins in C. botulinum isolates in homemade and commercial canned foods using the rapid, sensitive, and reliable Real-Time PCR technique. A total of 81 canned samples were collected, including 51 commercial and 30 homemade cans. The canned samples were initially enriched in a tryptone peptone glucose yeast (TPGY) medium. Subsequently, DNA isolation was performed using the foodproof® StarPrep Two Kit. After the isolation process, Real-Time PCR was conducted using the foodproof® C. botulinum Detection LyoKit and 5' Nuclease (Biotecon, R60240) kit. Among all samples, only 2 samples were positive. Both positive samples were found to contain neurotoxin type A, were commercially canned foods, and were packaged in glass containers. It was also determined that one was a mixed vegetable side dish and the other was roasted red pepper. Consequently, it was concluded that commercially available canned foods in Türkiye could still pose a potential risk to public health due to the presence of C. botulinum. Therefore, it was recommended to perform quality control analyses through food inspections in companies involved in canned food production and sales.*

Keywords: *Canned food, Clostridium botulinum, BoNT, Real-time PCR*

Received: February 16, 2024 *Accepted: June 5, 2024*

1. Introduction

Since *Clostridium botulinum* (*C. botulinum*) is a spore-forming bacterium that thrives in low-oxygen settings, canned foods provide the ideal conditions for it to flourish and produce toxins. When consumed, the bacterium's powerful neurotoxin can cause severe neurological symptoms, paralysis, and, in severe cases, respiratory failure. Botulism is a relatively rare neuroparalytic disease that poses a life-threatening risk to humans and animals. It is caused by botulinum neurotoxins (BoNTs) produced by organisms belonging to the *Clostridium* genus [1]. BoNTs are among the most potent toxins known, with lethal doses as low as 30-100 ng. Botulism, which develops through the consumption of food contaminated with neurotoxin, remains a serious type of food poisoning that can result in death [2]. The disease occurs when a food containing the toxin of one or more of the seven immunological types (A-

G) of *C. botulinum* is consumed [3]. It is known that botulism cases in humans are primarily caused by types A, B, E, and F [4].

In botulism cases, early symptoms typically include fatigue, weakness, nausea, vomiting, diarrhea, and dizziness. These symptoms are followed by manifestations such as blurred vision, mydriasis, ptosis, dry mouth, difficulty swallowing and speaking, bladder atony, and constipation. In progressive stages, the disease can lead to muscle weakness in the limbs, respiratory paralysis, and death [4, 5]. The responsible foods are usually products that undergo certain treatments for preservation, such as canning, pickling, and smoking, but where the elimination of *C. botulinum* spores is not guaranteed. When the preservation method applied to the food is inadequate, and it continues to be stored under conditions that allow the growth of *C. botulinum*, this contaminated food can harbor *C. botulinum* toxin, one of the deadliest toxins known to humans [6].

Canned foods have long been considered a convenient and popular choice for consumers due to their extended shelf life and ease of storage. However, concerns regarding the presence of *C. botulinum*, the bacterium responsible for causing botulism, have raised questions about the safety of these products. Botulism is a severe and potentially life-threatening form of food poisoning, primarily associated with improperly processed or stored canned goods.

In the early years of commercial canning, many problems related to *C. botulinum* existed, but intensive studies on processing procedures resulted in significant improvements in the 1920s and 1930s [7]. However, cases of foodborne botulism still occur today. The frequency of these cases varies depending on countries' food preferences, food preparation methods, food safety regulations, and even the available treatment options [8, 9]. Although botulism has been reported to arise from commercial canned foods, the majority of cases recorded since 1899 have been attributed to home-canned foods [10, 11]. However, since home-canned foods are often produced in rural areas and are not well-documented, the full extent of the problem associated with home canning is not known.

In developed countries, the presence and distribution of bacteria and toxins in high-risk foods are regularly monitored, and a reliable database is established to shed light on their development and support epidemiological studies, leading to intensified measures for control. However, it is notable that in Türkiye, comprehensive studies on this subject are limited, mainly consisting of case reports. A study by Güran and Öksüztepe [4] revealed that *C. botulinum* requires an acidic pH level, usually above 4.6, to produce toxins. According to initial research by Stumbo et al. [7], several *C. botulinum*-related concerns were seen in the early years of using tin cans for commercial canning.

The necessity to consider public health hazards beyond conventional canning methods is highlighted by the latest results in our study, which show positive samples in glass jars, demonstrating that the potential threat extends not only to canned items but also to canned products packaged in glass containers. The link between botulism cases and vegetable canning has been supported by a number of Turkish studies. According to Karsen et al. [2], green bean canning was a factor in the majority of Türkiye's botulism incidents between 1983 and 2017. In a comparable way, all botulism cases documented in Türkiye from 2017 to 2021 were linked to vegetable preserves, according to Bıcakcıoglu et al. [12], Daşkaya and Pek [5], and Botan et al. [8]. The precise type of neurotoxin that caused these cases, however, was not routinely recorded.

Between 2008 and 2019, 55 cases of foodborne botulism were reported in Iran by Soltan et al. [13]; 19 of these cases were connected to Type A neurotoxin. It is crucial to educate canning businesses and retailers about the possible risk of *C. botulinum* infection in canned foods to address this worrying issue. Rigorous quality control procedures must be put in place, including routine food inspections and adherence to strict hygiene standards throughout the production process. Additionally, teaching people about the proper handling, storage, and consumption of canned products helps reduce the number of instances of botulism. We intend to provide useful insights into the potential public health dangers associated with *C. botulinum* in this study by conducting it and building on previous literature. Studies conducted in other countries also support our findings. According to the Centers for Disease Control and Prevention [9], Type A neurotoxin was the primary cause of the majority of botulism occurrences in the United States during 2016 and 2017. Therefore, the aim of this study is to use the rapid, sensitive, and reliable Real-Time PCR technique to detect the presence of *C. botulinum* in homemade and commercial canned foods and to identify the A, B, E, and F toxin-producing genes in positive samples.

2. Materials And Methods

2.1. Sample collection

In this study, 51 commercial canned foods and 30 homemade canned foods were used as study material. The commercial canned foods were obtained from national supermarkets located in Şanlıurfa province, while the homemade canned foods were sourced from various regions of Türkiye. For commercial canned food collection, at least one sample was obtained from each brand sold in chain markets throughout Türkiye. Homemade canned foods were collected with at least one sample from regions where canned food production is predominant in Türkiye. The main criteria for selecting samples were that commercial canned foods had no expiration date when purchased, and that homemade canned foods did not have a production date of more than 6 months. Information regarding the collected samples is provided in Table 1.

Ethical Statement

The laboratory was used as material in the study, ethics committee permission is not required.

Table 1. Details of canned food samples

Type of Can	Can content	Number of Cans (pieces)	Can Material	City (Number of samples)
Commercial Cans	Garniture	4	Glass jar	Şanlıurfa
	Kidney bean	4	Tin	
	Roasted Eggplant	3	Glass jar	
	Boiled chickpeas	4	Tin	
	Tomatoes	4	Tin	
	Boiled fresh Bean	4	Tin	
	Okra	3	Glass jar	
	Mushroom	2	Glass jar	
	Roasted red pepper	3	Glass jar	
	Pea	5	Glass jar	
	Pea	3	Tin	
	Boiled dry bean	2	Tin	
	Sweetcorn	10	Tin	

Table 1. Continued

Type of Can	Can content	Number of Cans (pieces)	Can Material	City (Number of samples)
Home-made Cans	Mixed tomatoes and peppers	8	Glass jar	Elazığ (3) Malatya (1) Diyarbakır (1) Şanlıurfa (1) Gaziantep (1) Adıyaman (1)
	Tomatoes	8	Glass jar	Elazığ (3) Adana (1) Samsun (1) Şanlıurfa (1) Balıkesir (1) Adıyaman (1)
	Boiled fresh beans	5	Glass jar	Mardin (1) Sivas (2) Kayseri (2)
	Mixed Tomato Bean	5	Glass jar	Ankara (2) Elazığ (2) Erzurum (1)
	Mixed Roasted Eggplant and Capia Pepper	4	Glass jar	Şanlıurfa (2) Adıyaman (1) Mardin (1)
Total		81		

2.2. Enrichment of samples

Before performing PCR, an enrichment process was conducted on the canned samples. 1 mL of the homogenized canned samples was taken and added to Falcon tubes containing 9 mL tryptone peptone glucose yeast (TPGY) broth (Sigma-Aldrich). The tubes were then incubated under anaerobic conditions at 30°C for 96 hours. After the enrichment process, DNA isolation was performed.

2.3. DNA isolation

For DNA isolation, the manual method using the foodproof® StarPrep Two Kit - bulk purification (Biotecon, Catalog No. S40008) commercial kit was employed. Two hundred microliters (200 µL) of liquid or homogenized material from the pre-enriched canned samples were used for DNA isolation. Initially, the elution solution was heated to +70°C, and 200 µL of the sample material was added to a nuclease-free 1.5 mL microcentrifuge tube, along with 200 µL of Binding Buffer and 40 µL of Proteinase K. The mixture was incubated at +70°C for 10 minutes and then thoroughly mixed with the addition of 100 µL of isopropanol. Subsequently, the mixture was transferred to a filter collection tube. In the next step, the filter tube was centrifuged at 8,000 × g for 1 minute. After centrifugation, the bottom collection tube was discarded, and a new collection tube was placed under the filter tube. Five hundred microliters (500 µL) of inhibitor removal solution was added, and the tube was centrifuged at 8,000 × g for 1 minute. The discarded collection tube was replaced with a new one. The same process was repeated by adding 500 µL of wash solution and centrifuging at 8,000 × g for 1 minute. After repeating the wash step with 500 µL of wash solution, the tube was centrifuged at maximum speed for 10 seconds. The elution solution (previously heated) was then distributed to the samples in the microcentrifuge tube (200 µL). The tube was centrifuged at 8,000 × g for 1 minute. Subsequently, the filter tube was discarded, and the accumulated DNA was stored at -20°C in the collection tube.

2.4. Real-time PCR analysis

For the target bacterium *C. botulinum*, the foodproof® *C. botulinum* Detection LyoKit, 5' Nuclease (Biotecon, Catalog No. R60240) kit was utilized. This commercially available kit is capable of detecting various *Clostridium spp*., including *C. sporogenes*, *C. botulinum*, and *C. butyricum*, along with toxin genes A, B, E, and F (Table 2).

Table 2. Real-Time PCR mix (master mix)

Chemicals	**µL**
PCR-pure ddH2O	25 µL
foodproof® *C. botulinum* Detection LyoKit	25 µL
foodproof® *C. botulinum* Detection Control Template (PK)	25 µL
Isolated DNA	25 µL
Total	**100 µL**

The melt curves observed in the real-time PCR monitor are presented in Figure 1. The reaction mixture was prepared in 100 µL tubes, and the neurotoxins were detected using the Real-Time PCR instrument (Qiagen RotorGene Q, USA) (Table 3). The sample, positive control (PK), and negative controls (ddH2O- 25 µL) were prepared in the same manner.

Table 3. Real-Time PCR Protocol

Programs	**Denaturation**	**Cycle**	**Melt Curve**	**Cooling**
Analysis Mode	-	Quantitation Mode	Melt Curve Mode	-
Cycle	1	50	1	
Target (ºC)	95	95 60	95 37	40
Condition (ss:dd:sn)	00:04:00	00:05:00	00:00:05	00:00:50
Acceleration (°C/sn)	20	20	20	0.2
Reading Mode	-	Single	Continuous	--

2.5. Melting curve analysis (Tm)

Melting Curve analysis was employed to identify the A, B, E, and F neurotoxins (refer to Table 4). Fluorescence emission was monitored for 10 seconds at the end of each binding cycle. In case the sample tested positive, the amplified products underwent melting curve analysis (Tm). Denaturation was conducted at 95°C for 50 seconds, followed by annealing at 37°C for 20 seconds. Subsequently, the temperature was gradually increased at a rate of 0.2°C per second until reaching 85°C, while continuously collecting and recording fluorescence measurements. This process facilitated the determination of the presence of A, B, E, and F neurotoxins in the DNA of the positive samples.

3. Results and Discussion

The amplification curves (Cp values) of all samples (n=81), along with the corresponding positive and negative results are illustrated in Figure 1.

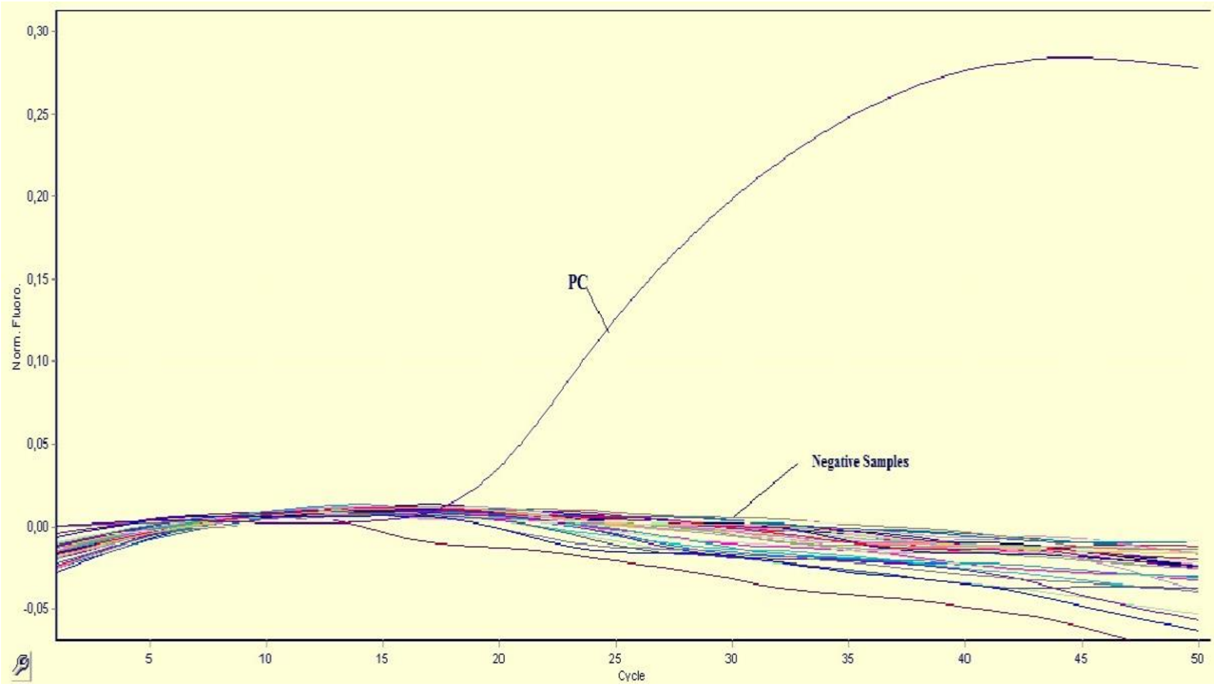

Figure 1. Re-run of negative samples and positive control (PC)

Figure 2 shows the amplification curves of the positive samples. Only two of the samples tested were positive for C. botulinum. The neurotoxin species of the positive samples was identified as BoNT A. One of these samples was commercial garnish canned and the other was a commercial roasted red pepper canned and in both cases, glass was the type of can used.

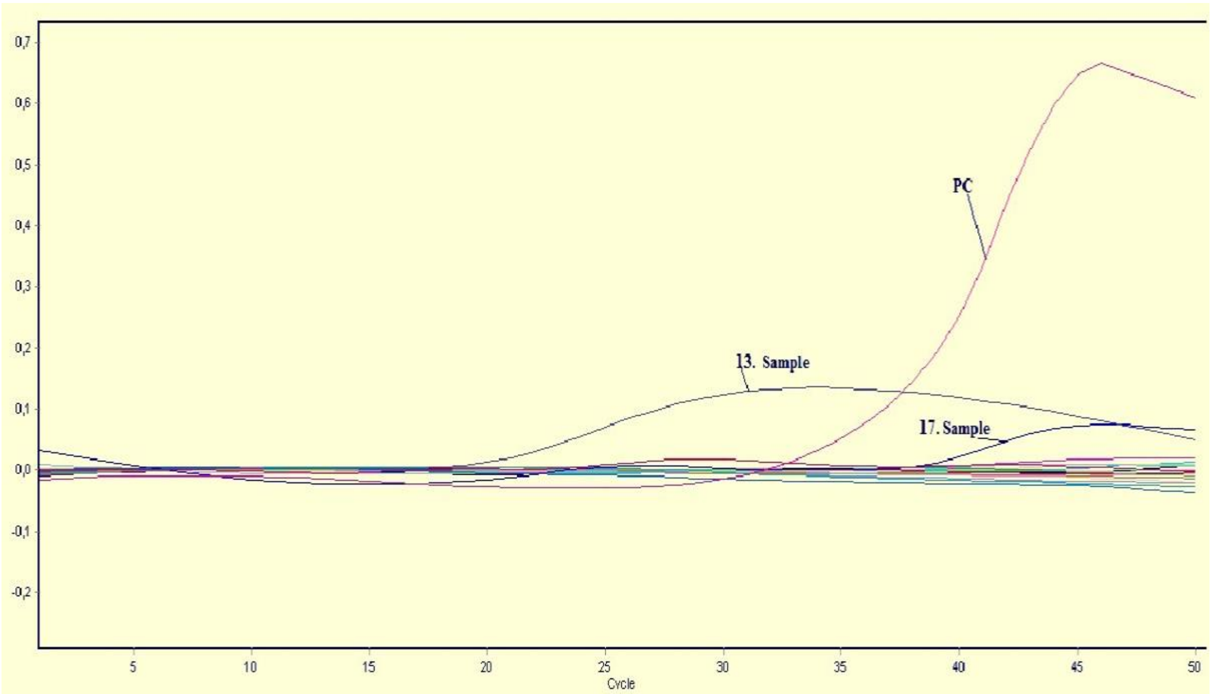

Figure 2. Positive samples and Positive control (PC)

The amplification curves of the Internal controls (IC) are presented in Figure 3 to demonstrate the accuracy of the PCR studies for all working samples.

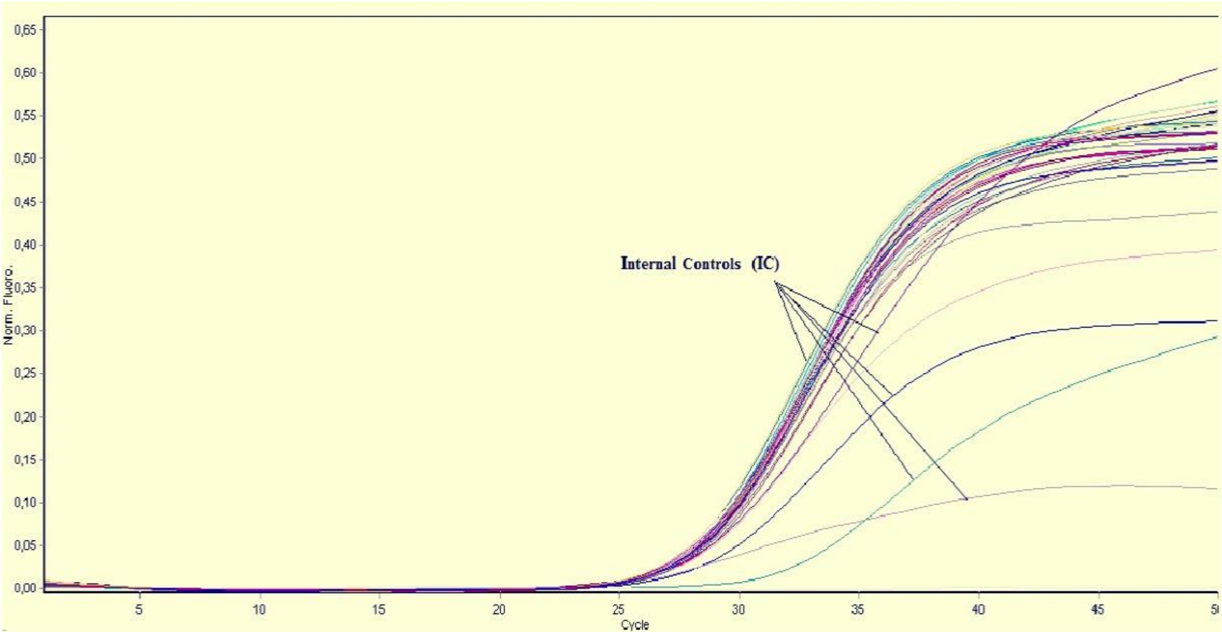

Figure 3. Internal controls (IC)

It was observed that all samples ran smoothly. Upon examining the results of the melting curve analysis, as intended, the melting curve of the positive control at 64.0°C ± 2.0°C in the FAM channel and 55.0°C ± 2.0°C and 62.5°C ± 2.0°C in the HEX channel is depicted in Figure 4 and Figure 5.

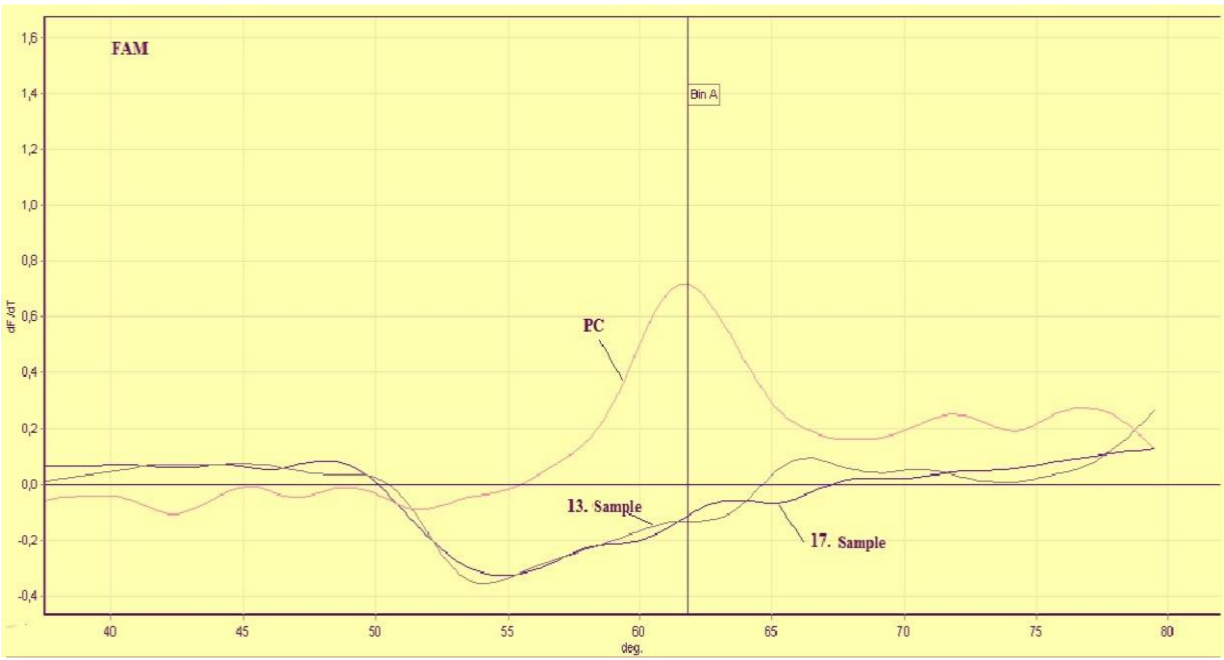

Figure 4. Melting curve analyzes of the FAM channel of positive control (64 °C ± 2 °C)

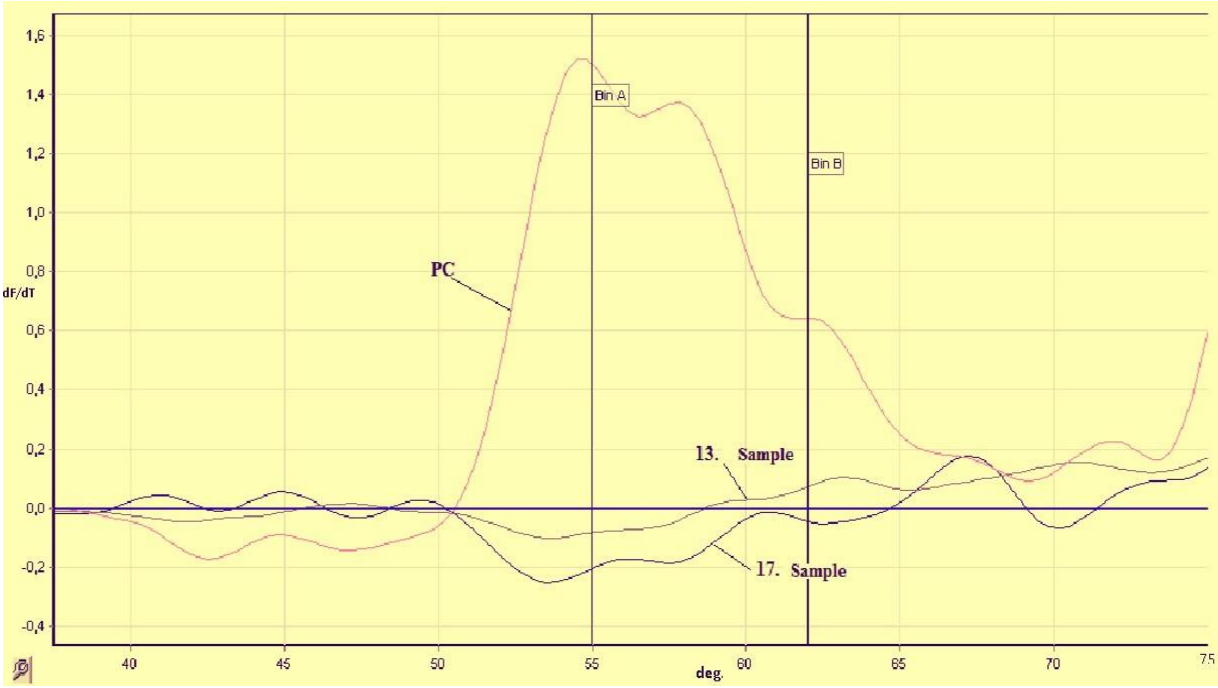

Figure 5. Melting curve analyzes of the HEX channel of positive control (55 °C ± 2 °C and 62.5 °C ± 2 °C)

Upon re-examining the positive samples, Real-time PCR analysis was conducted before and after enrichment, revealing no amplification in the samples before enrichment. After enrichment, both samples were found to be positive again (Figure 6).

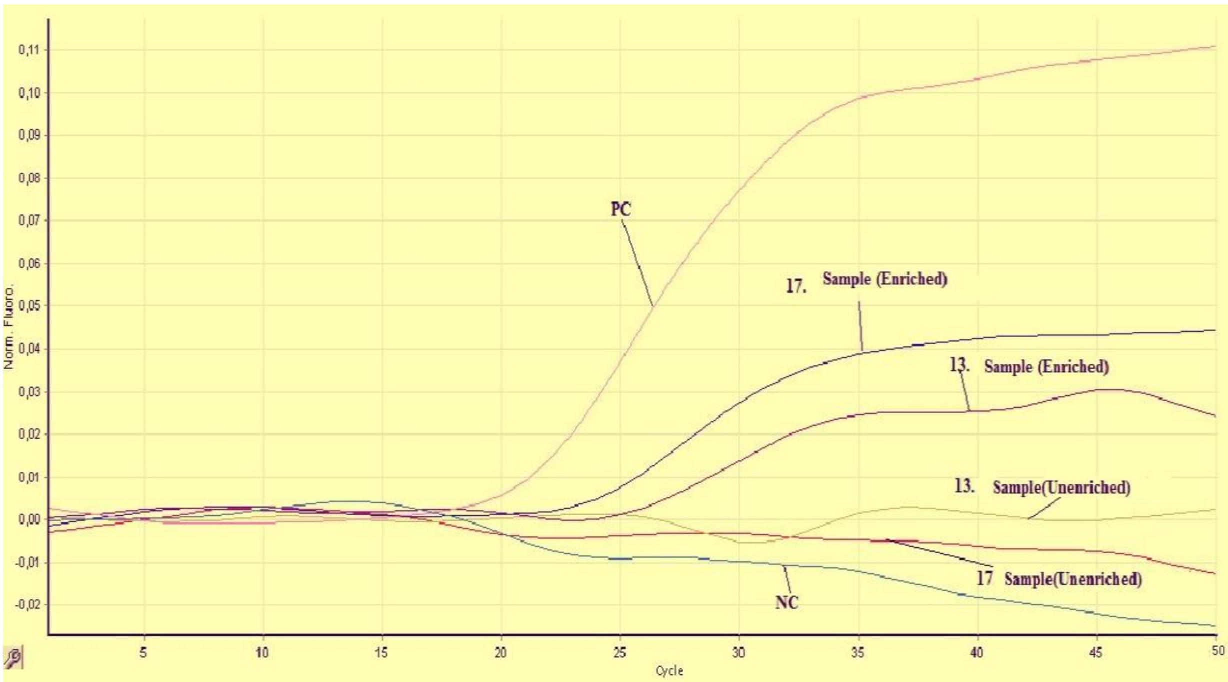

Figure 6. Real-time PCR results of positive samples (before and after enrichment)

C. botulinum spores are commonly found in nature. As a soil-borne bacterium, this pathogen naturally occurs in many fruits and vegetables. Canned foods made from contaminated fruits and vegetables play a significant role in the production and transmission of the toxin produced by this pathogen. To prevent human exposure to *C. botulinum* and ensure effective canning, it is necessary to prevent the contamination of raw foods with *C. botulinum* and to perform the canning processes properly [14].

Despite the parameters for canning being established and implemented many years ago, botulism cases are still reported. In 2014, the European Centre for Disease Prevention and Control reported 91 cases of botulism from 16 European Union countries, mostly in Romania, Poland, and Hungary [15]. In the United States, it was reported that 245 people were diagnosed with botulism in 2016, with 29 cases being foodborne; in 2017, 182 people were diagnosed with botulism, with 19 cases being foodborne; and in 2018, 231 people were diagnosed with botulism, with 18 cases being foodborne [9]. In Romania, between 2012 and 2018, 130 cases of foodborne botulism were reported [16]. In France, the annual incidence of botulism from 2010 to 2016 was reported as 17.4 cases per year, with foodborne botulism being the main cause [17]. In a hospital in Tehran, Iran, between 2008 and 2019, 61 cases of botulism were diagnosed, with 55 cases being foodborne [13]. Karsen et al. [2] reported 95 cases of foodborne botulism in Türkiye between 1983 and 2017, with 18 deaths. These findings support the idea that canned products still pose a public health risk for botulism.

In a study conducted in northern Italy between 2013 and 2020, 2187 food samples were collected, and the presence of *C. botulinum* toxin was investigated using multiplex real-time PCR. The toxin was detected in 16 samples, most of which were industrial products [18]. This study provides evidence that not only homemade canned foods but also commercial canned products may pose a risk to public health. Especially the detection of BoNT A genes in positive samples makes this situation even more risky. The presence of positive samples in commercial products is thought to be due to a failure to comply with the necessary parameters during the canning and storage stages. In contrast to our study, Davidson et al. [19] conducted a microbiological analysis of swollen cans and reported no detection of toxigenic strains. This discrepancy may be attributed to insufficient heat treatment of the products or post-processing contamination. The pathogen can withstand high temperatures and different environmental conditions and can cause intoxication through food [20]. It has been reported that consumption of commercial foods, which are the main cause of foodborne botulism, after the expiry date and storage at room temperature rather than refrigeration increase the risk of poisoning [11]. In the present study, although the positive samples were not problematic in terms of expiry date, the fact that they were stored at room temperature poses a potential risk to public health. This is because, depending on the storage conditions, the toxin level will increase until the canned food is consumed, causing more severe poisoning.

The absence of positive samples in homemade canned foods was surprising. This may be attributed to inadequate hermetic sealing during the production stage. Toxin production by bacteria requires the presence of a suitable anaerobic environment [21]. Additionally, the examined homemade canned foods in this study were mainly tomato-based, which may also contribute to the absence of positive samples. Tomato-based canned foods have low pH values, and pH is crucial for toxin production by the pathogen. The pH should be non-acidic (above 4.6) for the pathogen to produce the toxin [4]. For many acidic foods, including fermented and acidified vegetables, salsas, salad dressings, and others, keeping the pH below 4.6 is a critical control to prevent botulism [22]. It has been reported that there were many problems related to *C. botulinum* in the early years of commercial canned production using tin cans [7]. However, in our study, it is observed that both positive samples were found in glass jars. The findings of this study indicate that not only canned foods in tin cans but also those in glass jars may pose a risk to public health.

It is seen that one of the positive samples in the study was mixed vegetable canned food, and the other was roasted red pepper. Karsen et al. [2] reported that most cases of botulism in Türkiye between

1983 and 2017 were caused by green bean canned foods. This shows that canned vegetables sold and consumed in Turkey pose a risk for *C. botulinum.*

Considering the potential use of BoNTs as a biological weapon, foodborne botulism can still be considered an emerging risk disease. The detection of BoNT A genes in positive samples indicates that canned foods in Turkey still pose a risk to public health. However, heating can also inactivate BoNTs, depending on various factors such as medium composition, strain type, process temperature, etc. Different types of BoNT have different heat stability and some can even survive normal pasteurisation [23]. Rasooly and Do [24] reported that low-temperature pasteurisation of milk (63 °C, 30 min) inactivated BoNT of serotype A but not serotype B. Therefore, heat treatment of canned foods to inactivate BoNTs before consumption may reduce the risk in these products.

4. Conclusion

In conclusion, our study highlights the potential danger to public health posed by commercially available canned foods in Türkiye, as they may contain *C. botulinum*. This finding emphasizes the need for stringent quality control measures through regular food inspections in both canning companies and retail markets. It is crucial to prioritize effective hygiene practices and provide comprehensive education to producers and consumers regarding canning production procedures. It is crucial for canning businesses to implement and follow stringent hygiene measures throughout the whole production process in order to prevent food contamination and reduce the danger of foodborne botulism. This entails meticulous pH level monitoring, careful handling of chemicals, and extensive equipment cleanliness to prevent the growth of *C. botulinum*. Additionally, educating consumers about the possible dangers of canned foods and giving them instructions on how to handle, store, and consume them safely can make a big difference in reducing the number of cases of botulism. The right storage conditions, expiration dates, and the significance of checking cans for damage or bulging before consumption should all be explained to consumers. To ensure adherence to hygienic and safety standards, government regulatory bodies, food safety organizations, and public health authorities should work closely with canning businesses. To confirm that best practices are being followed and to quickly identify and address any possible problems, routine inspections and audits should be carried out. Prioritizing these actions will help canned food products be safer overall and of higher quality, reducing the risk of foodborne botulism and preserving the general public's health.

Ethical Statement

The laboratory was used as material in the study, ethics committee permission is not required.

Acknowledgments

This work was financially supported by the Harran University Scientific Research Projects Unit.

Compliance with the Research and Publication Ethics

This study was carried out in accordance with the rules of research and publication ethics.

Conflict of Interest

The authors declare no conflict of interest.

Financial Disclosure

Harran University Scientific Research Projects Unit

Author Contribution

H.D: Writing - Original draft preparation

A.Y., S.K.A., M.E.A: Methodology
A.Y., M.E.A: Formal analysis, Writing
S.K.A., M.E.A: Conceptualization, Methodology
H.D., S.K.A: Investigation
Resources All authors read and approved the final manuscript.

References

[1] Scalfaro, C., Auricchio, B., De Medici, D., Anniballi, F., "Foodborne botulism: an evolving public health challenge", *Infectious Diseases,* 51(2), 97-101, 2019. https://doi.org/10.1080/23744235.2018.1524584

[2] Karsen, H., Ceylan, M.R., Bayındır, H., Akdeniz, H., "Foodborne botulism in Turkey, 1983 to 2017", *Infectious Diseases*, 51(2), 91-96, 2019. https://doi.org/10.1080/23744235.2018.1524582

[3] Kanaan, M. H. G., Tarek, A. M., "Clostridium botulinum, a foodborne pathogen and its impact on public health", *Annals of Tropical Medicine and Public Health*, 23, 49-62, 2020. https://doi.org/10.36295/ASRO.2020.2357

[4] Güran, H.Ş., Öksüztepe, G., "Gıda kaynaklı botulizm ve önemi", *Fırat Üniversitesi Sağlık Bilimleri Veteriner Dergisi*, 26(3), 191-195, 2012.

[5] Daşkaya, H., İpek, H., "First case of botulism type A in Turkey", *Turkish Journal of Medical Sciences*, 49(4), 1349-1351, 2019.

[6] Rawson, A. M., Dempster, A. W., Humphreys, C. M., Minton, N. P., "Pathogenicity and virulence of *Clostridium botulinum", Virulence*, 14(1), 2205251, 2023. https://doi.org/10.1080/21505594.2023.2205251

[7] Stumbo, C. R., Gross, C. E., Vinton, C., "Bacteriological studies relating to thermal processing of canned meats", *Food Research*, 10, 260-272, 1945.

[8] Botan, Y., Öztürk, E., Şahin, O. T., Uğurlu, G., "An outbreak of foodborne botulism due to consumption of homemade roasted eggplant paste in Turkey", *Foodborne Pathogens and Disease*, 18(2), 118-121, 2021.

[9] CDC., "Botulism in the United States, 2016–2018", *Morbidity and Mortality Weekly Report*, 69(34), 1150-1154, 2020.

[10] Sofos, J. N., "Botulism in Home-Processed Foods", *In Food Poisoning, Routledge*, 171-203, 2019.

[11] Rasetti-Escargueil, C., Lemichez, E., Popoff, M. R. "Public health risk associated with botulism as foodborne zoonoses" *Toxins*, 12(1), 17, 2019. https://doi.org/10.3390/toxins12010017

[12] Bıcakcıoglu, M., Sırıken, B., Çelik, M., İnce, E., Ener, B., "Three clustered cases of botulism associated with canned green beans contaminated with *Clostridium botulinum" Foodborne Pathogens and Disease*, 13(2), 119-121, 2016.

[13] Soltan Dallal, M. M., Najar, F., Karbalaei Shabani, A., Ezoji, K., Shabani, M., Montazer Khorasan, M. R. "Investigation of the frequency of foodborne botulism in patients referred to Loghman Hospital in Tehran City, Iran, From 2008 to 2019", *International Journal of Medical Toxicology and Forensic Medicine*, 10(2), 28348, 2020. https://doi.org/10.32598/ijmtfm.v10i2.28348

[14] Lúquez, C., Edwards, L., Griffin, C., Sobel, J., "Foodborne botulism outbreaks in the United States, 2001–2017", *Frontiers in Microbiology*, 12, 713101, 2021. https://doi.org/10.3389/fmicb.2021.713101

[15] ECDC., "Annual epidemiological report 2016 – Botulism", *ECDE*, 2016. https://www.ecdc.europa.eu/en/publications-data/monitoring/all-annual-epidemiological-reports

[16] Negrut, N., et al., "Epidemiology of botulism in the north-western Romania 7-year survey", *Environmental Science and Pollution Research*, 28(45), 64234-64240, 2021. https://doi.org/10.1007/s11356-021-15576-w

[17] Rasetti-Escargueil, C., Lemichez, E., Popoff, M. R., "Human botulism in France, 1875–2016", *Toxins*, 12(5), 338, 2020. https://doi.org/10.3390/toxins12050338

[18] Benevenia, R., et al., "Foodborne botulism survey in Northern Italy from 2013 to 2020: Emerging risk or stable situation", *Food Control*, 132, 108520, 2021. https://doi.org/10.1016/j.foodcont.2021.108520

[19] Davidson, P. M., Pflug, I. J., Smith, G. M., "Microbiological analysis of food product in swelled cans of low-acid foods collected from supermarkets" *Journal of Food Protection*, 44(9), 686-691, 1981.

[20] Lake, D. E., Lesniewski, R. S., Anderson, J. E., Graves, R. R., Bremser, J. F., "Enumeration and isolation of mesophilic anaerobic sporeformers from cannery post-processing equipment", *Journal of Food Protection*, 48(9), 794-798, 1985. https://doi.org/10.4315/0362-028X-48.9.794

[21] Johnson, E. A., "*Clostridium botulinum*", *Food Microbiology: Fundamentals and Frontiers*, 487-512, 2019.

[22] Breidt, F., "BufferCapacity3 an interactive GUI program for modelling food ingredient buffering and pH", *SoftwareX*, 22, 101351, 2023. https://doi.org/10.1016/j.softx.2023.101351

[23] Munir, M. T., Mtimet, N., Guillier, L., Meurens, F., Fravalo, P., Federighi, M., Kooh, P., "Physical treatments to control Clostridium botulinum hazards in food", *Foods*, 12(8), 1580, 2023. https://doi.org/10.3390/foods12081580

[24] Rasooly, R., Do, P. M., "Clostridium botulinum neurotoxin type B is heat-stable in milk and not inactivated by pasteurization", *Journal of Agricultural and Food Chemistry*, 58(23), 12557-12561, 2010. https://doi.org/10.1021/jf1028398

Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Review Article

FERMENTED DAIRY PRODUCTS AND ORTHODONTICS

***Aybüke Asena ATASEVER İŞLER*[1]** ***Serap KILIÇ ALTUN*[2*]**

[1]Department of Orthodontics, Bolu Abant İzzet Baysal University, Faculty of Dentistry, Bolu, Türkiye
[2] Dep. of Food Hygiene and Technology, Harran University, Faculty of Veterinary Medicine, Şanlıurfa, Türkiye
Corresponding author: skilicaltun@harran.edu.tr

Abstract: *The oral environment constitutes a critical component of the human microbiome. The intake of probiotic products can exert beneficial effects on the oral microbiota, thereby reducing the risk of diseases. Nutrition is the most essential component not only for survival but also for maintaining a healthy lifestyle. This review aims to explore the impacts of probiotic products on oral flora and specifically address the role and significance of probiotics in the context of orthodontic treatment processes. For this purpose, two impartial reviewers examined the search results from four electronic databases (Web of Science, Google Scholar, Scopus, and Dergipark), and then selected articles that might be pertinent for full-text analysis. Data extraction and methodological flaws were addressed in articles that satisfied the inclusion criteria, and the quality of the studies was rated using the Cochrane Instrument for Risk of Bias Assessment. Fermented dairy products are acquired as a result of the fermentation of milk with the activity of beneficial and suitable microorganisms. In addition to fermentation bacteria, fermented milk products contain bacterial bioactive compounds and metabolites produced throughout fermentation. With this content, fermented dairy products are a unique composition due to the inclusion of ingredients that serve the ultimate product properties on the other side of nutrition. This review focuses on the available scientific data on the consumption of fermented milk products and their orthodontic utilities. Orthodontic patients avoid many types of food because their teeth are sensitive to chewing. Thus, these individuals consume significantly less protein and other essential nutrients, calcium, and some vitamins. This article provides an overview of the relationship between fermented milk products and orthodontic treatment. The role of fermented milk products in the nutrition of orthodontic patients is discussed to achieve a healthy mouth and general health.*

Keywords: *Nutrition, fermented dairy products, orthodontics*

Received: December 14, 2023 *Accepted: April 14, 2024*

1. Introduction

Nutrition has been adopted as the most basic component of a healthy lifestyle. To support healthy nutrition in societies, it is practiced to constantly explain healthy nutrition through health services [1]. In today's world, what we expect from food is not only to meet the energy required for our basic physical needs but also to contribute to human health. For this reason, food manufacturers are no longer looking for varieties that are both high-yielding and reducing exposure to disease and environmental impacts, not just to feed the ever-growing world population. Food suppliers are turning to products that will both meet the current consumer demand for healthy food products and provide health benefits. Because there are many expectations such as supporting the immune system, preventing chronic diseases, and providing enough energy without causing obesity [1, 2]. Considering all the necessary elements in nutrition, it is formed by determining the age, nutritional habits, likes and dislikes of individuals, etc.

[3, 4]. Milk is a good source of vital nutrients such as protein, enzymes, fat, lactose, vitamins, minerals, immunoglobulins, and hormones. Dairy products are an effective functional food not only to meet the nutritional needs of consumers but also to prevent many diseases such as osteoporosis, dental diseases, obesity, poor gastrointestinal health, cardiovascular diseases, and hypertension [4]. Studies on the effect of milk on oral health have shown that each component in the composition has many benefits. For example, it is known that glycomacropeptides in the composition of milk can prevent dental caries and gingivitis, and show antiviral and antibacterial effects. On the other hand, conjugated linoleic acid and lactoferrin protect oral health by modulating the immune system [5]. Fermented milk products have many properties such as anti-microbial, anti-mutagenic, anti-carcinogenic, and anti-hypertension and It is known that it has benefits on tooth mineralization metabolism. It is known that fermented milk products are the major carriers of probiotic microorganisms, and some clinical studies show the impacts of probiotic strains on dental health [3, 6-8]. However, there is a need for component-effect-based studies on this.

The success of orthodontic treatment hinges significantly on the patient's engagement and cooperation throughout the therapeutic process. Throughout this period, the challenges patients face while eating with orthodontic appliances become a matter of significant concern. A thorough understanding by orthodontists of how orthodontic devices affect their patients' dietary habits and nutritional regimens will enable the provision of more effective nutritional strategies and guidance [4, 5]. This comprehension is crucial in ensuring that orthodontic interventions are not only mechanically successful but also supportive of maintaining overall patient health and well-being during treatment.

Orthodontic problems in individuals can occur as soon as they are born with the wrong habits they have acquired. There is a bilateral relationship between nutrition and orthodontics. Treatment where orthodontic treatment affects food intake and nutritional quality affects the speed of orthodontic treatment [3]. At the same time, it is accepted that individuals who receive orthodontic treatment are in the high-risk group for dental caries and periodontal diseases, and for this reason, the foods that individuals take with diet are important [9]. On the other hand, orthodontic devices have been used for a long time to treat malocclusions and weak jaw relations, but their influences on patient food intake showed that total fat, copper, manganese, and lipid levels from food were significantly reduced [10]. Nutritional follow-up of patients is very important considering the effects of nutrients on periodontal status and the effects of orthodontic treatment on nutrition [6]. Because diet determines oral flora, periodontal condition, and healing. Orthodontic treatment is related to many factors, from the growth and development of the craniofacial complex to the movement of teeth through periodontal tissues [6]. Today, the increasing importance of oral aesthetics; directs individuals to orthodontic treatment in order to improve their dental appearance. With the development of fixed orthodontic treatment applications, the demand for orthodontic treatment has increased. Bands, brackets, and wires used in orthodontic treatment create areas suitable for the retention of nutrients. Thus, it becomes difficult to maintain oral hygiene and an ecological area is created for microorganisms that cause tooth decay. There are many studies in the literature reporting an increased incidence of caries during fixed orthodontic treatment [8]. *Streptococcus mutans* and lactobacilli are the leading microorganisms that cause dental caries [11]. Fermented dairy products can play a significant role in the diet, both to strengthen the anatomy in orthodontic treatments and to prevent caries that may occur during treatment. Another positive aspect of fermented dairy products in the diet of orthodontic patients is that they are a unique food choice for braces wearers because most fermented milk products are soft and require little chewing [12]. This study aims to evaluate the future of orthodontic treatment by considering the role of fermented milk products in reducing the bacteria that cause dental caries.

2. Materials and Methods

Two impartial reviewers examined the search results from four electronic databases (Web of Science, Google Scholar, Scopus, and Dergipark), and then selected articles that might be pertinent for full-text analysis. Data extraction and methodological flaws were addressed in articles that satisfied the inclusion criteria, and the quality of the studies was rated using the Cochrane Instrument for Risk of Bias Assessment.

3. Results and Discussion

3.1. Fermented Milk Products

Fermented milk products have shown a significant increase in consumption in the last years and market tendency suggests that this will increase even more. As fermented dairy products contribute to healthy living and increase life expectancy, there is an increasing consumer interest owing to the health benefits these products offer [13]. Fermentation processes often increase the bioavailability and nutritional relevance of many foods. The fermentation effect of lactic acid bacteria leads to the conversion of antinutritional agents such as galactose and lactose in milk, increasing bioavailability [14]. The conversion of lactose to lactic acid is the major phenomenon in fermentation. The enzymes of lactic acid bacteria change milk carbohydrates into oligosaccharides, some of which have prebiotic features [15]. The formation of other compounds produced by lactic acid bacteria depends on the composition of the raw milk used, the microbial strains, and the status of the fermentation process. The most known types of lactic acid bacteria utilized for milk fermentation are *Bifidobacteria* or *Streptococcus thermophilus*, which is usually associated with *Lactobacilli* [15, 16].

3.2. Probiotics and Orthodontics

The term probiotic is derived from the words “pro” and “biota” and means “vital, vital”. While defining the concept of probiotics, Elie Metchnikoff (17), who won the Nobel Prize in 1908, pointed out that harmless live bacteria in the human body could have beneficial effects on the host, and then in 1998, Guarner and Schaafsman defined probiotics as “live microorganisms that provide significant health gains beyond supporting a healthy life”. According to the definitions of the World Health Organization and the Food and Agriculture Organization, probiotics are live microorganisms that provide health benefits to the host when taken in sufficient amounts [18, 19]. The bacteria most frequently used as probiotics are *lactobacilli* and *bifidobacteria*, as seen in Table 1. *Lactobacilli*, one of the lactic acid bacteria, are bacteria found in large quantities, especially in fermented animal foods. Adhesion of probiotics with pathogenic bacteria in the oral microbiota It creates a promising alternative in terms of oral and dental health as they compete for the surface and inhibit them with their metabolites. The usability of probiotic cultures in oral health studies has gained importance in recent years. Today, products such as chewing gum, toothpaste, lozenges, tablets, and mouthwash containing oral probiotics produced in some countries such as America and Canada are offered to people in the markets. In a study by Jose et al.(2013), in which they compared the effects of systemic probiotic paste consumption and topical probiotic toothpaste application on Streptococcus mutans levels in the plaques of orthodontic patients, they reported that probiotic toothpaste was more effective than systemic consumption [20]. Several studies have indicated the effectiveness of various *Lactobacillus* species in orthodontic treatment. These include *Lactobacillus reuteri* strain [21], *Lactobacillus paracasei* strain [22], as well as *Lactobacillus salivarius*, *Lactobacillus plantarum*, and *Lactobacillus rhamnosus* [23]. Additionally, research suggests that *Lactobacillus brevis* may offer benefits due to its anti-inflammatory properties [24, 25, 26]. *Streptococcus salivarius* M18 strain has been predominantly assessed for its anti-caries activity [27]. *Bifidobacterium* is another species that has shown promise in positively impacting periodontal disease [25]. In a separate study, the *Streptococcus salivarius* K12 strain was evaluated for its potential to influence oral malodor parameters [26]. In theory, multi-strain probiotic products may offer synergistic and symbiotic benefits due to interactions among their different strains. However, there is limited evidence suggesting that probiotic strains could also inhibit each other. For instance, while the production of hydrogen peroxide and bacteriocins may effectively inhibit endogenous strains such as Streptococcus mutans, they might also inadvertently suppress other probiotic strains within the same formulation, potentially reducing the product's overall efficacy [28]. For probiotics to exert their intended effects, bacteria must adhere to and colonize the inner surface of the mouth [28]. Effective probiotic activity cannot be expected if there is biofilm present on the oral surface or if the oral pH is not conducive to bacterial survival [22]. In such scenarios, it may be necessary to either increase the concentration of the probiotic or extend the duration of administration to observe potential clinical improvements in the patient [29]. Notably, changes in bacterial composition were observed only after six weeks of administering an oral probiotic preparation [30].

Table 1. Widely used probiotics in the dairy industry [16,19]

Genus	Probiotic strain	Some products containing probiotics
Lactobacillus	*L. plantarum, L. acidophilus, L. paracasei, L. rhamnosus, L. casei, L. gasseri, L. crispatus, L. bulgaricus, L. reuteri*	yogurt, kefir, buttermilk, cheese,
Lactococcus	*L. rhamnosus, L. lactis, L. plantarum, L. reuteri, L. curvatus, L. casei, L. acidophilus*	yogurt, cultured butter, buttermilk, cheese, sour cream
Bifidobacterium	*B. breve, B. longum, B. Bifidum, B. catenulatum, B. animalis, B.infantis*	yogurt, kefir, fermented milk drinks, fermented dairy products for infants and children, probiotic supplements
Streptococcus	*S. thermophilus, S. salivarius, S. sanguis, S. oralis, S. mitis*	yogurt, cheese, fermented milk drinks, probiotic supplements, oral health products such as probiotic lozenges or chewing gums
Propionibacterium	*P. freudenreichii, P. jensenii*	Swiss cheese, Emmental cheese, Gruyère cheese
Bacillus	*B. subtilis, B. coagulans, B. laterosporus, B.cereus*	traditional dairy products such as yogurt, cheese, or fermented milk drinks
Pediococcus	*P. pentosaceus, P. acidilactici*	cheese, fermented milk drinks, yoğurt, buttermilk, probiotic supplements
Saccharomyces	*S. boulardii, S. pastorianus, S. cerevisiae*	kefir, fermented dairy drinks, probiotic supplements, cheese
Akkermansia	*A. muciniphila*	it is not currently a common component of dairy products on the market.
Escherichia	*E. coli Nissle 1917*	supplement form, pharmaceutical preparations
Enterococcus	*E. faecium, E.munditii, E.durans, E. hirae*	cheese, fermented dairy products, probiotic supplements,
Leuconostoc	*L. mesenteroides*	cheese, fermented milk drinks, buttermilk, sour cream, probiotic supplements

Probiotic microorganisms should be able to produce antimicrobial compounds, stimulate the immune response, not transfer the antibacterial resistance feature to pathogenic microorganisms, and should not be pathogenic. It was reported that the first anaerobic bacteria isolated from the oral cavity was the *Bifidobacteria* group [31]. The effects of probiotic bacteria occur through three mechanisms; reducing the number of pathogenic bacteria, improving the immune system, and altering the metabolism (enzymatic activity) [32]. Probiotics also exert individual and strain-specific effects. Probiotics are transient microorganisms on the host, but they must be continuously colonized in order to have efficacy. This does not pose a disadvantage since they do not need to be colonized. The use of probiotics in oral health focuses on maintaining a balanced oral microbiome, which may help prevent conditions such as dental caries (tooth decay), gingivitis (gum inflammation), and periodontal disease (gum disease).

Some probiotic strains commonly associated with oral health include *Lactobacillus* species: Certain strains of *Lactobacillus*, such as *Lactobacillus reuteri* and *Lactobacillus acidophilus*, have been studied for their potential benefits in promoting oral health by reducing levels of harmful bacteria in the mouth and supporting a healthy balance of the oral microbiome [6, 11].*Streptococcus salivarius*: This bacterium is a natural inhabitant of the oral cavity and is known to produce bacteriocins, which are proteins that can inhibit the growth of other bacteria, including pathogenic strains. Some strains of *Streptococcus salivarius* have been investigated for their potential to promote oral health [9, 12].*Bifidobacterium* species: Certain strains of *Bifidobacterium* have also been studied for their potential benefits in oral health, although research in this area is still relatively limited compared to other probiotic strains [11, 12]. While these probiotic strains may have potential benefits for oral health, including supporting orthodontic treatment by reducing the risk of oral health problems, more research is needed to fully understand their effectiveness and optimal use in this context. In a study on saliva and plaque, probiotics have been shown to recover one week after discontinuation of treatments [7].

3.3. Prebiotics and Orthodontics

Prebiotics are food components consisting of indigestible short-chain carbohydrates that selectively induce the activity and growth of probiotics [33, 34]. They are hydrolyzed by probiotics in the large intestine and transformed into beneficial molecules instead of being digested in the upper portion of the digestive tract [35]. What is expected from prebiotics is that they pass through the upper parts of the digestive system and have the ability to increase the growth of beneficial bacteria in the large intestine. Another role of prebiotics is to increase the absorption and synthesis of B vitamins [36]. As a result, in order to prebiotics to be utilized by the bacteria in the large intestine, they must be able to resist acid hydrolysis in the stomach, transit into the large intestine undamaged, or be absorbed in the small intestine [37]. Galactooligosaccharides, fructooligosaccharides, and inulin are some naturally occuring prebiotics in foods. Lactosucrose, lactulose, isomaltooligasaccharide, glucooligosaccharides, xylooligosaccharides are generally synthesized prebiotics. Fermented dairy products comprise probiotics, which become accepted as synbiotics as prebiotics are added they support enhancing calcium absorption [38]. Many studies have demonstrated a relationship between bone density and various clinical problems in dentistry. Chugh et al. (2013) reviewed that information on bone density in an area of the oral cavity may be helpful in assessing orthodontic tooth movement as well as in planning to improve the success rate of treatment [39]. Children fed milk and orange juice containing inulin and oligofructose for a year had increased bone calcium homeostasis; their bone mineral density increased by 45% and their bone mineral content increased by 15%, according to previous studies [40, 41].

4. Conclusions

The impact of orthodontic appliances on patients' nutritional functions has been well-documented within the scientific literature. Developing a comprehensive understanding of the challenges faced during alimentation with these appliances will enable the creation of more effective patient education materials. Consequently, this could lead to an increase in patient compliance, and thereby, the overall success of the treatment can be targeted for improvement. Quantitative and qualitative research methods could be employed to systematically measure these difficulties and to examine the effects on the dietary quality of patients with orthodontic appliances. Armed with this data, tailored dietary recommendations could be developed to minimize the adverse effects of orthodontic appliances on daily dietary habits, integrating these recommendations into the treatment process to optimize therapeutic outcomes.

Food choice restriction was encountered by individuals wearing various types of orthodontic appliances. The causes of these dietary limitations were most commonly attributed to three factors: The impact of the appliance on the teeth, such as the physical/functional challenges associated with wearing an appliance which necessitates substituting hard or crunchy foods (for example, carrots, hard chocolate) with softer alternatives (such as pasta, yogurt); the dietary restrictions recommended by orthodontists aimed at reducing decalcification risks (for instance, avoiding sweets, fizzy drinks) and avoiding breakages (for example, not chewing gum, toffees) as well as the apprehension concerning potential damage to the appliance. The wearers of orthodontic appliances often impose dietary restrictions upon themselves to mitigate the embarrassment or inconvenience caused by food particles becoming lodged in their appliances, such as meats or green vegetables.

There is a bilateral relationship between nutrition and orthodontic treatment; the delivery of orthodontic treatment affects food intake, and the quality of nutrition affects the speed of orthodontic treatment. A proper diet strengthens orthodontic treatment by providing all the elements necessary to keep oral tissues healthy and help bone remodel. On the other hand, ensuring the maximum comfort of the patient while applying orthodontic treatment affects the diet at a minimum level. Fermented dairy products are important food groups in orthodontic treatment with their content and consistency.

In conclusion, this review highlights the potential orthodontic benefits associated with the consumption of fermented dairy products. Fermented dairy products, such as yogurt and cheese, are not only rich in essential nutrients but also contain beneficial microorganisms, bacterial bioactive compounds, and metabolites produced during fermentation. These components contribute to the unique composition of fermented dairy products and offer potential health benefits beyond basic nutrition.

Orthodontic patients, who often experience sensitivity and dietary restrictions, may benefit from incorporating fermented dairy products into their diet. These products provide essential nutrients such as protein, calcium, and vitamins while offering probiotic properties that support oral health. Furthermore, fermented dairy products are often soft and require minimal chewing, making them suitable choices for individuals undergoing orthodontic treatment.

Probiotic microorganisms found in fermented dairy products, including lactobacilli and bifidobacteria, have been studied for their potential to reduce the risk of dental caries, gingivitis, and periodontal disease. Additionally, prebiotics found in some fermented dairy products, such as inulin and oligofructose, may support bone health and calcium absorption, which is particularly relevant for orthodontic patients.

Overall, the consumption of fermented dairy products represents a promising avenue for supporting oral health and general well-being during orthodontic treatment. Further research and clinical studies are warranted to explore the specific mechanisms and optimal strategies for integrating these products into orthodontic care effectively. By emphasizing the role of nutrition, including fermented dairy products, in orthodontic treatment, this review contributes to a comprehensive understanding of holistic approaches to oral health management.

In conclusion, probiotic dairy products emerge as a promising option for enhancing oral health. Future research should expand to include the effects of other dairy products, such as ice cream and various types of cheese, on oral health beyond dental caries and periodontal disease. Additionally, detailed surveys will be necessary to evaluate changes in the profile of research conducted post-pandemic, the relationship between the flow of publications and the countries most affected, and the impact of the use of probiotics in dentistry.

Conflicts of interest

All authors have none to declare.

Ethical statements

The authors declare that this document does not require ethics committee approval or any special permission. This review does not cause any harm to the environment and does not involve the use of animal or human subjects.

Authors' contributions

AA-Aİ and S-KA made a substantial contribution to the literature search and article write-up. Both authors read and approved the final document.

References

[1] Crowley, J., Ball, L., & Hiddink, G. J., “Nutrition in medical education: a systematic review”, *The Lancet Planetary Health*, 3, 379-389, 2019. Doi: 10.1016/s2542-5196(19)30171-8

[2] Hsieh, P. Y. H., & Ofori, J. A., “Innovations in food technology for health”, *Asia Pacific Journal of Clinical Nutrition*, 16, 65-73, 2007.

[3] Can, T., Aksoy, A., & Baygut, H., “Fixed Orthodontic Treatment and Nutrition”, *Journal of Int Dent Sci*, 6, 25-33, 2020.

[4] Demirgül, F., & Sağdıç, O., “The Effect of Fermented Milk Products on Human Health”, *European Journal of Science and Technology*, 13, 45-53, 2018. Doi: 10.31590/ejosat.377798

[5] Chandan, R. C., “Nutritive and health attributes of dairy ingredients”, *Dairy Ingredients for Food Processing*, 387-419, 2011. Doi: 10.1002/9780470959169.ch16

[6] Singh, N., Tripathi, T., Rai, P., & Gupta, P., “Nutrition and Orthodontics-Interdependence and Interrelationship”, *Res Rev J Dent Sci*, 5, 18-22, 2017.

[7] Ravn, I., Dige, I., Meyer, R. L., & Nyvad, B., “Colonization of the oral cavity by probiotic bacteria”, *Caries Res*, 46, 107-112, 2012. Doi: 10.1159/000336960

[8] Lara-Carrillo, E., Montiel-Bastida, N. M., Sanchez-Perez, L., & Jorge Alanis-Tavira, J., “Factors correlated with developing caries during orthodontic treatment: Changes in saliva and behavioral risks”, *Journal of Dental Sciences*, 7, 218-224, 2012.

[9] Bani, M., & Ayhan Bani, A., “Oral Health Awareness and Rehabilitation Oral Health of Orthodontics Patients”, *J Dent Fac Ata Uni*, 24, 442-451, 2014.

[10] Miresmaili, A., “Orthodontic Appliance Effect on Nutrition”, *Nutrition Bytes*, 20(1), 2016.

[11] Alp, S., Comparison of The Effect of Systemic Use of Kefir, Containing Probiotic Bacteria, and Local Administration of Probiotic Toothpaste on Microbial Colonization in the Saliva in Patients Undergoing Orthodontic Treatment, Master thesis of Selcuk University, 19-68, 2017.

[12] Khatri, J. M., & Kolhe, V. D., “Nutrition and orthodontics”. *International Journal of Orthodontic Rehabilitation*, 9, 163, 2018. Doi: 10.4103/ijor.ijor_21_18

[13] Bourrie, B. C., Willing, B. P., & Cotter, P. D., “The microbiota and health promoting characteristics of the fermented beverage kefir”, *Frontiers in microbiology*, 7, 647, 2016. Doi: 10.3389/fmicb.2016.00647

[14] Shiby, V. K., & Mishra, H. N., “Fermented milks and milk products as functional foods—A review”, *Critical Reviews in Food Science and Nutrition*, 53, 482-496, 2013. Doi: 10.1080/10408398.2010.547398

[15] Granier, A., Goulet, O., & Hoarau, C., “Fermentation Products: Immunological Effects on Human and Animal Models”, *Pediatric Research*, 74, 238-244, 2013.

[16] Fijan, S., “Microorganisms with claimed probiotic properties: an overview of recent literature”, *International Journal of Environmental Research and Public Health*, 11(5), 4745-4767, 2014. Doi: 10.3390/ijerph110504745

[17] Dugas, B., Mercenier, A., Lenoir-Winjkoop, I., Arnaud, C., Dugas, N., & Postaire, E., “Immunity and Probiotics”, *Immunology Today*, 20, 387-390, 1999.

[18] Klaenhammer, T. R., “Probiotic bacteria: today and tomorrow”, *J Nutr*, 130, 415-416, 2000.

[19] Kerry, R. G., Patra, J. K., Gouda, S., Park, Y., Shin, H. S., & Das, G., “Benefaction of probiotics for human health: A review”, *Journal of Food and Drug Analysis*, 26, 927-939, 2018. Doi: 10.1016/j.jfda.2018.01.002

[20] Rotimi, V. O., & Duerden, B. I., “The development of the bacterial flora in normal neonates”, *J Med Microbiol*, 14, 51-62, 1981.

[21] Şahin, T., & Özmeriç, N., “Drink Called Enjoyment: The Role of Kefir in Dentistry and Its Future in Periodontology”, *Hacettepe University Faculty of Health Sciences Journal*, 8, 438-452, 2021. Doi:10.21020/husbfd.815116

[22] Panitantum, V., “The story of probiotics, prebiotics & synbiotics”, A seminar presentation at Kasetsart University, Bangkok, under the auspices of BIOTEC, National Science and Technology Development Agency, Bangkok, Thailand, January 26, 157-161, 2004.

[23] Imran, M. Y. M., Reehana, N., Muralitharan, G., Thajuddin, N., & Dhanasekaran, D., “New Formulations and Products in Prebiotic Food”, In Advances in Probiotics (s. 437-447, Academic Press, 2021.

[24] Vernazza, C. L., Rabiu, B. A., & Gibson, G. R., Human colonic microbiology and the role of dietary intervention: Introduction to prebiotic. In G. R. Gibson & R. A. Rastall (Eds.), Prebiotics: Development & Application (s. 1-28), John Wiley & Sons, Ltd., West Sussex, England, 2006.

[25] Allen, J. M., Jaggers, R. M., Solden, L. M., Loman, B. R., Davies, R. H., Mackos, A. R., & Bailey, M. T., “Dietary oligosaccharides attenuate stress-induced disruptions in immune reactivity and microbial B-vitamin metabolism”, *Frontiers in Immunology*, 10, 1774, 2019. Doi: 10.3389/fimmu.2019.01774

[26] Gibson, G. R., “Prebiotic”, *Journal Gastroenterology Supplement*, 18, 287-298, 2004.

[27] Thammarutwasik, P., Hongpattarakere, T., Chantachum, S., Kijroongrojana, K., Itharat, A., Reanmongkol, W., & Ooraikul, B., “Prebiotics-A Review”, *Songklanakarin Journal of Science & Technology”*, 31, 401-408, 2009.

[28] Chugh, T., Jain, A. K., Jaiswal, R. K., Mehrotra, P., & Mehrotra, R., "Bone density and its importance in orthodontics", *Journal of Oral Biology and Craniofacial Research,* 3, 92-97, 2013. Doi: 10.1016/j.jobcr.2013.01.001

[29] Silawanichnakul, W., "New trend of functional beverages", *Food Focus Thailand,* 2, 26-28, 2007.

[30] Khan, A., Dhar, A., Upadhyay, V., Munawar, T. M., & DV, S. P., "A Review on Prebiotics and Probiotics", *International Journal of Plant Biotechnology*, *9*(1), 29-32, 2023.

[31] Vivekananda, M. R., Vandana, K. L., & Bhat, K. G., "Effect of the probiotic Lactobacilli reuteri (Prodentis) in the management of periodontal disease: a preliminary randomized clinical trial", *Journal of Oral Microbiology,* 2(1), 5344, 2010.

[32] Sookkhee, S., Chulasiri, M., & Prachyabrued, W., "Lactic acid bacteria from healthy oral cavity of Thai volunteers: inhibition of oral pathogens" *Journal of Applied Microbiology*, 90(2), 172-179, 2001.

[33] Shimauchi, H., Mayanagi, G., Nakaya, S., Minamibuchi, M., Ito, Y., Yamaki, K., & Hirata, H., "Improvement of periodontal condition by probiotics with Lactobacillus salivarius WB21: a randomized, double-blind, placebo-controlled study", *Journal of Clinical Periodontology,* 35(10), 897-905, 2008.

[34] Riccia, D. D., Bizzini, F., Perilli, M. G., Polimeni, A., Trinchieri, V., Amicosante, G., & Cifone, M. G., "Anti-inflammatory effects of Lactobacillus brevis (CD2) on periodontal disease" *Oral Diseases,* 13(4), 376-385, 2007.

[35] Hojo, K., Mizoguchi, C., Taketomo, N., Ohshima, T., Gomi, K., Arai, T., & Maeda, N., "Distribution of salivary Lactobacillus and Bifidobacterium species in periodontal health and disease", *Bioscience, Biotechnology, and Biochemistry,* 71(1), 152-157, 2007.

[36] Burton, J. P., Chilcott, C. N., Moore, C. J., Speiser, G., & Tagg, J. R., "A preliminary study of the effect of probiotic Streptococcus salivarius K12 on oral malodour parameters", *Journal of Applied Microbiology*, 100(4), 754-764, 2006.

[37] Di Pierro, F., Zanvit, A., Nobili, P., Risso, P., & Fornaini, C., "Cariogram outcome after 90 days of oral treatment with Streptococcus salivarius M18 in children at high risk for dental caries: results of a randomized, controlled study", *Clinical, Cosmetic and Investigational Dentistry,* 107-113, 2015.

[38] Jose, J. E., Padmanabhan, S., & Chitharanjan, A. B., "Systemic consumption of probiotic curd and use of probiotic toothpaste to reduce Streptococcus mutans in plaque around orthodontic brackets", *American Journal of Orthodontics and Dentofacial Orthopedics*, 144(1), 67-72, 2013.

[39] Jivraj, F. E., "Effectiveness of the Lorodent Probiotic Lozenge in Reducing Plaque and S. Mutans Levels in Orthodontic Patients: a Randomized, Double-Blind, Placebo-Controlled Trial" University of Toronto (Canada), 2015

[40] Habib, S., "Assessment of the therapeutic potential of a dental probiotic in orthodontic patients affected by gingivitis: a randomized control trial", *University of Toronto (Canada),* 2016.

[41] Montero, E., Iniesta, M., Rodrigo, M., Marín, M. J., Figuero, E., Herrera, D., & Sanz, M., "Clinical and microbiological effects of the adjunctive use of probiotics in the treatment of gingivitis: A randomized controlled clinical trial", *Journal of Clinical Reriodontology,* 44(7), 708-716, 2017.