Conference Paper/Konferans Bildirisi

GAZİ
JOURNAL OF ENGINEERING SCIENCES

Detection of Personality Features From Handwriting By Machine Learning Methods

Hilal Müsevitoğlu[*a], Ali Öztürk[b, c], Fatiha Nur Başünal[d]

Submitted: 25.01.2023 Revised: 23.05.2023 Accepted: 16.06.2023 doi:10.30855/gmbd.0705064

Keywords: Handwriting analysis, Machine learning, Graphology

[a,*] HAVELSAN,
06510 - Ankara, Türkiye
Orcid: 0000-0001-5791-790X
e mail: hmusevitoglu@havelsan.com.tr

[b] HAVELSAN,
06510 - Ankara, Türkiye
Orcid: 0000-0002-1797-2039

[c] KTO Karatay University,
Computer Engineering Department.
42020 - Konya, Türkiye
Orcid: 0000-0002-1797-2039

[d] HAVELSAN,
06510 – ANKARA, Türkiye
Orcid: 0000-0002-0896-1194

[*]Corresponding author:
hmusevitoglu@havelsan.com.tr

ABSTRACT

Handwriting contains a lot of information about the person writing it and is a sign of personality traits represented by neurological patterns in the brain. In other words, our brain and subconscious actually shape our character as a result of our habits. It is possible to get an idea about the mood of the individual by examining the handwriting. Joy, sadness, anger and anxiety are some of them. In this study, handwritings of people belonging to different professions and age groups were collected. Feature extraction methods was applied on these articles by applying word and line detection, slant, pressure, page layout and similar image processing methods. The obtained features formed the inputs of the dataset. Personality traits such as calm, optimistic, emotional, extrovert, which were estimated using graphology, were added to the dataset as outputs. Then, this dataset was applied to Random Forest (RF), Naive Bayes (NB), Decision Tree, Support Vector Machines (SVM), Logistic Regression, Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) algorithms. According to the performance metrics used, the Random Forest algorithm gave the most successful results in terms of accuracy, precision and f1-score metrics. For this algorithm, the accuracy, precision, recall and f1 score values were found to be 0.90, 0.91, 0.84 and 0.85, respectively. Furthermore, the results of the personality analysis were compared with the results of the personality test performed by the expert psychologist. As a result of this comparison, it was seen that there was a 73% match.

Makine Öğrenmesi Yöntemleri ile El Yazısından Kişilik Özelliklerinin Tespiti

ÖZ

El yazısı, yazan kişi hakkında birçok bilgiyi barındırır ve beyindeki nörolojik desenler tarafından temsil edilen kişilik özelliklerinin işaretidir. Diğer bir deyişle beynimiz ve bilinçaltımız aslında alışkanlıklarımızın bir sonucu olarak karakterimizi biçimlendirmektedir. El yazısı incelenerek bireyin içinde bulunduğu ruh hali hakkında bir fikre varmak mümkündür. Sevinç, hüzün, öfke ve kaygı bunlardan bazılarıdır. Bu çalışmada farklı meslek ve yaş gruplarına ait kişilerin el yazıları toplanmıştır. Bu yazılar üzerinde kelime ve satır tesbiti, eğim, bası, sayfa boşluklarının bulunması ve benzeri görüntü işleme yöntemleri uygulanarak özellik çıkarımı yapılmıştır. Elde edilen özellikler veri setinin giriş sütunlarını oluşturmuştur. Grafoloji kullanılarak tahmin edilen sakin, iyimser, duygusal, dışa dönük gibi kişilik özellikleri ise veri setine çıkış sütunları olarak eklenmiştir. Daha sonra bu veri seti Rastgele Orman, Naive Bayes, Karar Ağacı, Destek Vektor Makinaları, Lojistik Regresyon, XGBoost ve LightGBM algoritmalarına uygulanmıştır. Doğruluk, kesinlik ve f1-score performans metriklerine göre en başarılı sonucu Rasgele Orman algoritması vermiştir. Bu algoritma için doğruluk, kesinlik, hatırlama ve f1 skor değerleri sırası ile 0.90, 0.91, 0.84 and 0.85 olarak bulunmuştur. Ayrıca kişilik analizi sonuçları, uzman psikolog tarafından yapılan kişilik testi sonuçları ile karşılaştırılmıştır. Bu karşılaştırma sonucunda %73 oranında eşleşme olduğu görülmüştür.

Anahtar Kelimeler: El yazısı analizi, Makine öğrenmesi, Grafoloji

To cite this article: H. Müsevitoğlu, A. Öztürk, F. N. Başünal, "Detection of Personality Features From Handwriting By Machine Learning Methods," *Gazi Journal of Engineering Sciences,* vol. 9, no. 2, pp. 200-212, 2023. doi:10.30855/gmbd.0705064

1. Introduction

Writing, which is a part of our lives, not only expresses what is written on paper, but also contains clues about our feelings. In this period when technology is advancing rapidly, handwriting still maintains its importance and existence, although it seems that handwriting has been replaced by the internet, media and e-mail in an undeniable extent. Everyone uses handwriting as well as technology in social or business life. Traces, patterns, shapes, symbols left from the moment the tip of the pen touches the paper have the ability to reflect a person's identity in a unique way, just like a fingerprint or DNA. About a hundred years ago, German Professor W. Preyer said that writing was created by the brain, not by training certain muscles [1]. Subsequent studies have also revealed that the most influential organ in the article is the brain. From this point of view, it is seen that handwriting is an effective tool for directly accessing the human brain and therefore psychology. Therefore, character analysis from handwriting has the potential to be applied in many different fields. Looking at the literature; forensic investigations, human resources (recruitment process), psychological counseling and guidance services, training etc. are used in the fields. Studies in which handwriting and personality analysis are associated are generally examined under the science of Graphology. Graphology is a field of study that covers inferences about a person's personality and character, based on a person's handwriting [2]. In order to make inferences in graphology, methods such as writing inclination, pen pressure, characteristic of certain letters, spaces between lines and words, inclination and size of letters are used.

There are different studies in the literature using machine learning algorithms for personality analysis. In the study carried out by Champa and Kumar; MATLAB was used to process features such as slope, pen pressure, characteristic of letters ‘y’ and ‘t’ [3]. In another study on determining personality traits from handwriting using convolutional neural networks (CNN); it has been proposed to determine the personality, structure and symbol trait analysis of an individual from the handwritten image. CNN, which is based on symbol analysis, was used as a classification method. The margin, the spaces between the lines, the spaces between the words, the printing and the oblique or inclination and certain letter feature which were not done in previous researches were examined [4]. Santana et. al. estimated identity analysis from handwriting using 5 different graphological parameters as input on a dataset with 29 participants, namely "vertical position", "combination of letters", "pressure force", "thinning area" and "letter a". They used Artificial neural networks (ANN) and Support Vector Machine algorithms. As a result of the study, 99.34% success was achieved from the dataset consisting of 29 authors. Later, when the dataset was expanded to 70, the success rate decreased to 92% [5]. In a different study [6], the identity of the author was extracted from the writing style. A new feature based on Graphology or Forensic techniques, which is digitized and classified by a decision fusion block by neural networks, is presented. Offline analysis was done using handwritten images. A success rate of 94.6% was obtained on the dataset consisting of 100 samples. Another study discussed the latest developments in computer graphology systems. In addition to the features such as page margin, baseline, line and word spacing, line direction, inclination of words and letters, size of letters, pen pressure, writing speed, the letter ‘t’, letter ‘i’, letter ‘f’ and the letter ‘y’ can also be effective in writing analysis. It has been mentioned that different rule-based analysis methods such SVM and artificial neural networks (ANN) are frequently used in graphology [7]. Anand et. al. used graphology, aptitude testing and personality testing to help people get acquainted with various career fields and choose the appropriate profession. The study considered the fundamental aspects of human behavior and analysis. Analyzes used were Aptitude Test, Psychometric Test (Myers-Briggs Type Indicator) and handwriting. From the handwriting samples taken, features such as spacing between words, left margin, right margin, font pressure, letter inclination, letter size, and the size of the letter "I" were calculated. The integration of all three modules offers individuals suitable career options [8]. In another study based on machine learning approach for handwriting personality analysis and author identification, a baseline, spacing between words, left margin, letter pitch, bar height in letter “t” attributes were used to predict an individual's personality. The polygonization method was used to extract features from the base. For this purpose, pattern matching was used for the letter “t” and one line was used for other features (projection profiles of the text). With the help of these features, the identity of the author was determined by the developed tool. The IAM database and Python programming language were used to develop the tool [9].

The handwriting features used in our study are; the right, left, top and bottom margins of the page, lines and spaces between words, pen pressure and slant of the text. These were chosen, because they reflect the personality analysis characteristics which are most commonly used in graphology. The machine learning algorithms used are; NB Classifier, Decision Trees, RF, SVM, Logistic Regression, XGBoost and LightGBM. Python language and OpenCV Library are used for the implementation of algorithms and feature extraction, respectively.

2. Materials and Methods

This section introduces the dataset used in the study, the personality test applied to the people from whom the handwritings were taken and the machine learning algorithms.

2.1. Dataset

The dataset used in this study was created by collecting handwriting samples from people of different ages and working groups. A ready-made text was given to the people and they were asked to write the same text on a blank A4 paper in their own handwriting.

Handwriting samples were taken from a total of 70 people. The dataset to be applied to machine learning algorithms was obtained by extracting the features from the handwriting samples with image processing methods. Section 3 Methods Used for Feature Extraction describes the feature extraction process in detail. The values in the dataset were normalized to the range of 0-1 by applying the min-max normalization method.

2.2. Personality test

The Adjective-Based Personality Test used in the study was developed by Bacanlı et al. in line with the Five Factor Personality theory [10]. The scale consists of 40 items and can be completed in 10-15 minutes by the participants. In practice; users are expected to mark the ones closest to them from the 40 adjectives given as two opposite poles. The choices are (1) Very appropriate, (2) Fairly appropriate, (3) Somewhat appropriate, (4) Neither appropriate, nor appropriate, (5) Somewhat appropriate, (6) Fairly appropriate, and (7) Very appropriate. The original Adjective-Based Personality Test consists of five dimensions: (1) Neuroticism-7 items, (2) Extraversion-9 items, (3) Openness to Development-8 items, (4) Agreeableness-9 items, and (5) Responsibility-7 items.

In this study, in the light of the personality analysis obtained from the handwritings, the items were correlated separately by a specialist psychologist and evaluated by taking the average values of the items, depending on whether they were suitable for personality types or not. Finally, the correlation of the acquired personality types with the expert assessment was examined.

2.3. Machine learning algorithms

2.3.1. Naive Bayes classifier

It is a simple method based on Bayes' theorem. It gives good results especially when the number of features in the dataset is high. Therefore, it is widely used in the classification of texts. Despite its simplicity, NB classifier can achieve comparable performance with complex classification methods such as decision tree and neural network classifier. NB explicitly calculates probabilities for classifier hypotheses. In this respect, it has an advantage over algorithms that do not use probability. During training, each example contributes to the probability of the hypothesis. Naive Bayes classifier is based on Bayesian formula given in Equation 1.

$$(A|B) = P(B|A)P(A)/P(B) \tag{1}$$

Where,

P(B | A) : The probability that B is true given A.
P(A | B) : Given B, the probability that A is true.
P(A) : The probability that A is true.
P(B) : The probability that B is true.

2.3.2. Decision trees

Decision trees search the hypothesis space using intersection combinations and try to reach the hypothesis that best fits the dataset. The expressive power of the hypotheses expressed as an intersection combination is high. The first classification application of decision trees was with the ID3 algorithm [11]. Later, the version C4.5, which can also handle continuous values, has emerged. One of the advantages of decision trees is that the tree formed as a result of the training is understandable by people. The internal structure of other machine learning algorithms is incomprehensible to humans. While creating the decision tree, the entropy (2) and information gain (3) are used in the selection of the next node.

$$Entropy(S) = -\sum_{i=1}^{c} p_i \, log_2(p_i) \tag{2}$$

$$Gain(S, A) = Entropy(S) \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{3}$$

We used "gini" index (4) for the measure of split quality in the implementation of Decision Tree.

2.3.3. Random forest

The random forest algorithm is based on the logic of creating small decision trees where the output of all of them are combined, rather than a single large decision tree. Thus, a more powerful classifier is obtained with the ensemble logic. The algorithm used to generate small decision trees in a random forest is the same as for normal decision trees. The difference in random forest is how the dataset is used when creating the trees. First, a separate dataset is created by randomly selecting the rows in the dataset. The same row may appear in a new set more than once. In addition, a few randomly selected features are used in the new dataset, not all features.

Thus, training of each tree in the forest is done with a new dataset randomly selected from the original dataset and also with randomly selected features from the features in the new set. In this way, several small trees are formed, varying with the size of the dataset and the number of features. Since the bootstrapping of new datasets is done by random selection, some rows may not be selected at all during this process. This is called out-of-bag. These lines are then used to measure the performance of the random forest. During the creation of small decision trees in the random forest, the Gini Index (4) criterion is generally used for the selection of features to be placed in the nodes [12].

$$Gini = 1 - \sum_{i=1}^{C} (p_i)^2 \tag{4}$$

2.3.4. Support vector machines

The original version of the support vector machines algorithm provides linear classification of data belonging to two different classes. In this algorithm, there is the concept of margin, which is used to separate the samples in the dataset from each other. Linear classifiers such as Perceptron classify data by separating them with a single linear line [13]. Therefore, this method provides a rough classification and is likely to fail to classify new incoming data. In the SVM method, on the other hand, there is a margin instead of a line separating the data from one another. Margin is the name given to the area between two parallel lines. The SVM algorithm has to keep the classification error to a minimum while maximizing this margin. This is a classic optimization problem. In Figure 1, the margin is shown in yellow. There are data belonging to the classes on two parallel border lines indicated by the dashed line. These points are called support vectors. The SVM method takes its name from these vectors.

2.3.5. Logistic regression

Logistic regression is a machine learning algorithm used in binary classification problems (when the target is categorical). Logistic regression basically uses a logistic function defined in the function to model a binary output variable [15]. The purpose of Logistic Regression is to discover the link between the features and the probability of a particular outcome. In the logistic function equation (5), x is the input variable.

$$Logistics\ function = \frac{1}{1+e^{-x}} \tag{5}$$

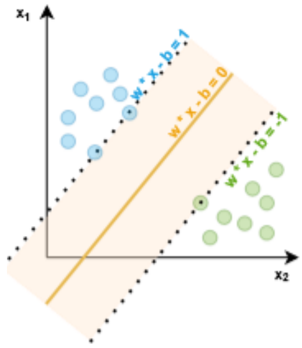

Figure 1. Representation of margin and support vectors for two-dimensional space [14]

2.3.6. XGBoost

Boosting means combining several weak learners to obtain a stronger learner. The boosting methods basically train predictors sequentially and add them to an ensemble, each one trying to correct its predecessor. The gradient boosting method trains the new predictor using the residual errors of the previous one and uses gradient descent over an objective function to minimize prediction errors in the next model. Gradient Boosting Decision Trees (GBDT) method train multiple shallow decision trees and the final prediction is obtained by integrating all tree predictions.

XGBoost algorithm uses Classification and Regression Tree (CART) as the elementary tree. The algorithm improved the optimization method of the cost function on traditional GBDT, and introduced a regular term to speed up training and reduce overfitting. XGBoost also includes instance and feature sampling as well as parallel histogram algorithm.

The objective function of XGBoost for K trees can be defined as

$$F_{obj} = \sum_{i=1}^{N} Cost(y_i, \hat{y}_i) + \sum_{i=1}^{K} \Omega(f_i) \tag{6}$$

Where the first term is usually the mean squared error, while the second term is

$$\Omega(f_i) = \gamma T + \frac{1}{2}\lambda|\omega|^2 \tag{7}$$

Where T denotes the number of leaves in each tree and γ is used to penalize T to avoid overfitting. The second term provides $L2$ regularization where ω is the weight scores of the leaves and λ is used to punish ω to improve the generalization performance.

For the *t-th* iteration, the F_{obj} objective function can be written according to Taylor series expansion and approximated as

$$F_{obj}^{t} = \sum_{i=1}^{N} Cost\left[y_i, \hat{y}_i^{(t-1)} + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t) \tag{8}$$

Where, $g_i = \partial_{\hat{y}_i^{(t-1)}} Cost\left(y_i, \hat{y}_i^{(t-1)}\right)$ is the first order gradient and $h_i = \partial^2_{\hat{y}_i^{(t-1)}} Cost\left(y_i, \hat{y}_i^{(t-1)}\right)$ is the second order gradient of the cost function. By means of h_i, the objective function can quickly converge to the optimal value [16].

2.3.7. LightGBM

LightGBM is a decision tree based emsemble learning method which uses the histogram algorithm to find the best split point and leaf-wise strategy for tree growing with depth limit [17]. In Figure 2, the growth of the tree by depth is shown. It uses the same basic principles of XGBoost like bagging, early stopping, regularization, multiple loss functions, etc. Unlike XGBoost, LightGBM does not use sorted features to find the opt imal split points. Instead of this, it utilizes an optimized histogram-based algorithm which improves training efficiency and memory usage. During training, the histogram-based algorithm buckets the continuous feature values into discrete bins and the feature histograms are constructued using these bins.

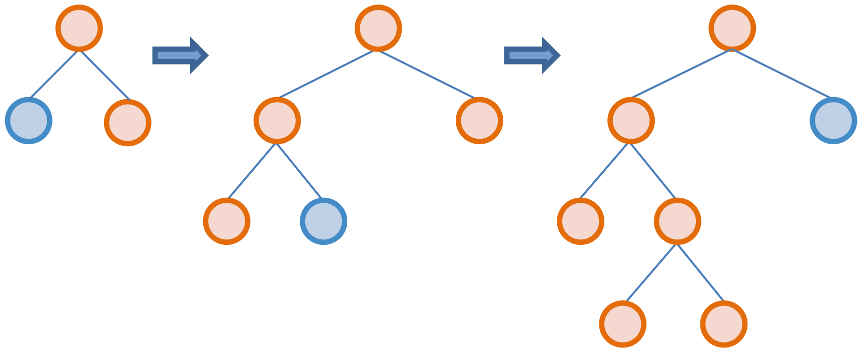

Figure 2. Leaf-wise tree growth

Since the number of bins is much smaller than the number of samples, histogram building increases the computational complexity. To remedy this situation, LightGBM use two new techniques called Gradient-Based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) to reduce both the number of samples and features, respectively. LightGBM uses GOSS to keep the data instances with larger gradients while performing random sampling on the instances having small gradients. This instance-based weighting strategy for training next model is also used by Adaboost algorithm. However, XGBoost uses residual errors of the previous tree for the next model instead of instance weights. The other technique which makes LightGBM more efficient is EFB. This technique merges exclusive features into a single feature thus achieving dimensionality reduction to improve efficiency in histogram building.

The other novelty in LightGBM is that the trees grow in leaf-wise manner with depth limitation instead of the traditional level-wise growth. The leaves with the greatest split gain at each iteration is selected. Since this strategy may easily cause overfitting by growing the tree much larger, the depth of tree must be set in LightGBM to improve accuracy and avoid overfitting [18]. As a result, LightGBM uses less memory, reduces training time of the XGBoost model, and improves accuracy. It is well suited for dealing with problems involving large amounts of data [19].

3. Methods Used For Feature Extraction

The steps applied for features extraction from handwritten images are given below.

1. Load the hand writing image
2. Locate the word frames
3. Calculate the average of the spaces between the word frames
4. Determine the lines
5. Calculate the average spaces between the lines
6. Determine the left margin from the left most pixel of the left most word frame
7. Determine the right margin from the right most pixel of the right most word frame
8. Determine the top margin from the top most pixel of the top most word frame
9. Determine the bottom margin from the bottom most pixel of the bottom most word frame
10. Calculate the layout of the page using the left, right, top and bottom margins
11. Determine the pen pressure using the average of the gray level values
12. Determine the slope of the text using the minimum area rectangle method

For personality analysis, inferences were made on the handwriting samples by using the features expressed as items below.

3.1. Right, left, top and bottom margins of the page

Rectangles drawn around key components such as words and letters in the handwritten image were used to calculate the margins of the page. Figure 2 shows the plotting of the right, left, upper and lower margins.

- Left margin: The left edge coordinate of the leftmost rectangle of the resulting rectangles
- Right margin: The rightmost coordinate from the right edges of the resulting rectangles
- Upper margin: Coordinate where the upper edge of the resulting rectangles is lowest
- Bottom margin: Coordinate where the bottom edge of the resulting rectangles is highest

Then, the following steps determines the personality:

- If the width on the right side of the page is too large; lively and active, if the width is small; serious and cautious personality
- If the left part is too wide; fearful of risk, if the width is small; it shows an impatient personality.
- If the width is too high at the top of the page; cautious, if the width is small; an impatient personality
- If the width is too wide at the bottom, thin in emotions and behaviors, if the width is less; it shows a nervous, high economic anxiety personality.

3.2. General layout

Equation 6 is used to check if the page is in a certain layout.

$$Page\ Layout = \left(\frac{(LM + RM)}{2} + \frac{(UM + LoM)}{2}\right)/2 \qquad (9)$$

Where, LM is Left Margin, RM is Right Margin, UM is Upper Margin and LoM is Lower Margin.
Based on the formula, the average spacing on the page margins means that:

- If it is less than 10%; shy, introverted familial personality
- If it is between 10% and 25%; balanced and harmonious personality
- If it is more than 25%; Shows generous, extroverted personality.

According to the rectangles drawn around the basic components such as words and letters in Figure 3; plotting the right, left, upper and lower margins is shown.

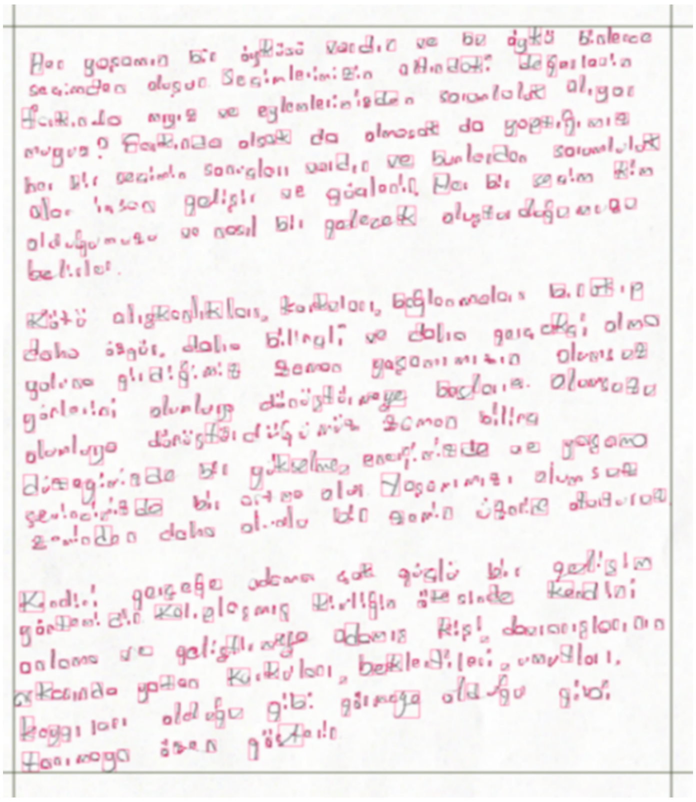

Figure 3. Indication of margin lines

3.3. The amount of space between words

After the words were identified in the handwriting, the spaces between the words and their averages were found. In Figure 4, the detected words and their numbering according to the lines are shown. If there is too much space between words; then it means introverted, fond of freedom person. If the space is less; it shows a sociable and active personality.

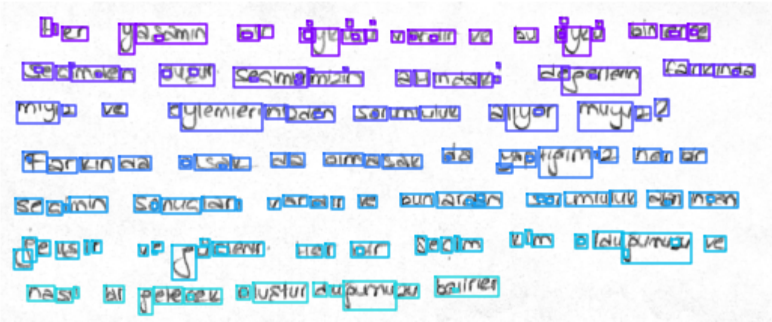

Figure 4. Indication of spaces between words

3.4. The amount of space between lines

After the words were identified, the lines formed by the words were found as shown in Figure 5. The spaces between these lines represent the spaces between the lines. Therefore, the average spaces between the lines were calculated.

If there is too much space between the lines; the tendency of the person to look from a calm and broad perspective is high. If it is less; this indicates a personality that loves movement and crowds.

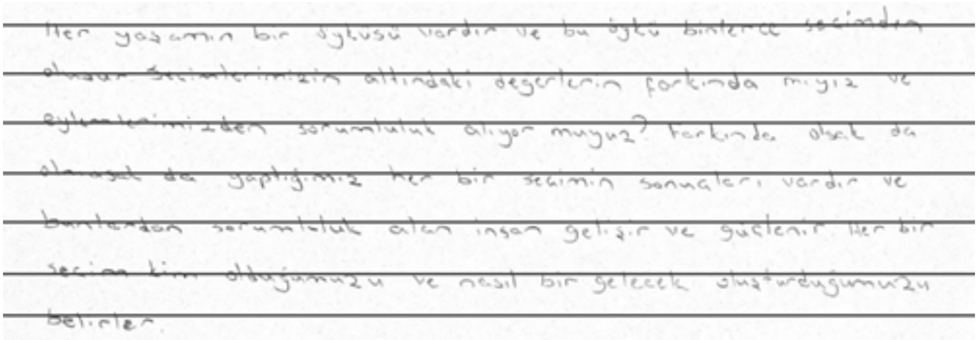

Figure 5. Indication of spaces between lines

3.5. Pressure applied to the pen

The average of the gray level tones in the text was found by calculating the closeness to black. If the pressure applied to the pen is too much; the person is stressed, tense and irritable. If it is low; this shows a fragile and graceful personality.

3.6. Slope of text

In order to find the slope of the entire text, the angle that the text makes with respect to the y-axis is calculated. If the slope of the text is downwards, it show pessimistic personality. Upward slant indicates an optimistic, straight, balanced personality.

4. Experimental Results

The first 10 rows of the dataset obtained by calculating all input characteristics for each person are shown in Table 1. Therefore, there were 13 input features in the dataset which were obtained by applying the relevant Open CV library functions to handwritten images as discussed in Section 3.

Table 1. Input properties of the dataset acquired

Slope	Press	Left Margin	Right Margin	Upper Margin	Lower Margin	Word Spacing	Line Spacing	Page Layout
2.2	247	266	146	147	1316	20	179	469
1.8	247	77	155	174	214	42	465	155
1.9	249	588	345	742	3120	54	106	1199
0.6	245	229	22	731	1817	13	154	700
0,6	245	53	111	42	800	1	32	252
2	244	62	0	25	166	53	79	63
1.2	248	693	0	466	2220	38	179	845
12.4	249	0	146	0	2418	3	97	641
0.5	250	549	254	426	3363	35	105	1148

Table 2 gives the personality traits suggested by the graphology. These were used as the outputs of the machine learning algorithms depending on the input characteristics given in Table 1.

Table 2. Output values obtained from personality analysis

Optimistic	Balanced	Stressful	Impatient	Leery	Uneasy	Emotional	Introvert	Extrovert	Calm

Therefore, the dataset was created by combining the personality traits shown in Table 2 to the corresponding inputs of Table 1. The threshold values of the input properties given in Table 3 were used to determine the personality outputs. The thresholds are subjective choices representing the general characteristics of the collected handwritings. For example; the slope threshold was chosen by considering the handwritings on which the lines are apperantly slant. The press threshold was chosen by looking at the handwritings which were written by applying relatively higher pressure than the others and so forth. Although, the threshold values are subjective, they can be used in classification of new handwriting samples. The assumption is that the same steps proposed in this study will be used to determine the personality from the new handwritings.

Table 3. The threshold values and personality outputs

Input feature	Threshold value	Personality output
Slope	>1.5	Optimistic
Press	<248.8	Stressful
Left Margin	<88	Impatient
Right Margin	>100	Balanced
Upper Margin	>342	Leery
Lower Margin	<875	Uneasy
Lower Margin	>875	Emotional
Word Spacing	>20	Introvert
Line Spacing	>171	Calm
Page Layout	>1000	Extrovert

The performance results of the machine learning algorithms were obtained using the parameters given in Table 4. During our experiments, these parameters were observed as the ones which mostly effected the performance of the algorithms. The values of the parameters given for each algorithm provided the highest scores of performance metrics for the respective algorithm. Some of the parameters in Table 4 have default values already provided by the Python implementation. In our evaluation, we changed the values of these ones to other options or values. However, we observed that the performance of the algorithms decreased in those cases. Therefore, the parameters default values retained. There are other parameters of the algorithms as well, but we did not touch the remaining default parameters of the algorithms.

Table 4. Parameters used in machine learning algorithms

Algorithms	Parameters used
Naive Bayes	priors=None, var_smoothing=1e-5
Support Vector Machine	C=5, cache size=200, coef0=0.02, decision_function_shape=ovr, degree=3, gamma=scale, kernel=poly, max_iter=-1, probability=False, random_state=1, shrinking=True, tol=0.001
Logistic Regression	C=10, fit_intercept=True, intercept_scaling=1, max_iter=300, multi_class=auto, penalty=None, random_state=1, solver=lbfgs, tol=0.0001
Decision Tree	Ccp_alpha=0.07, criterion=gini, min_samples_leaf=1, min_samples_split=3, random_state=1, splitter=best
Random Forest	bootstrap=True, ccp_alpha': 0.03, criterion=gini, min_samples_leaf=1, min_samples_split=2, n_estimators=100, oob_score=False, random_state=1, warm_start=False
LightGBM	boosting type=dart, colsample bytree=1.0, learning_rate=0.15, max_depth=-1, min_child_samples=10, min_child_weight=0.01, min_split_gain=0.12, n_estimators=80, n jobs=-1, num_leaves=10, reg_alpha= 0.003, reg_lambda=0.003, subsample=0.8, subsample_for_bin=200
XGBoost	booster=gbdt, gamma=1, enable_categorical=False, learning_rate=0.12, min _child_weight=0.01, n_estimators=80

The 3-fold cross-validation method was applied to the machine learning algorithms. Thus, performance metrics were obtained separately for each personality trait. The comparison of the machine learning algorithms according to the average of the performance metrics is given in Figure 6.

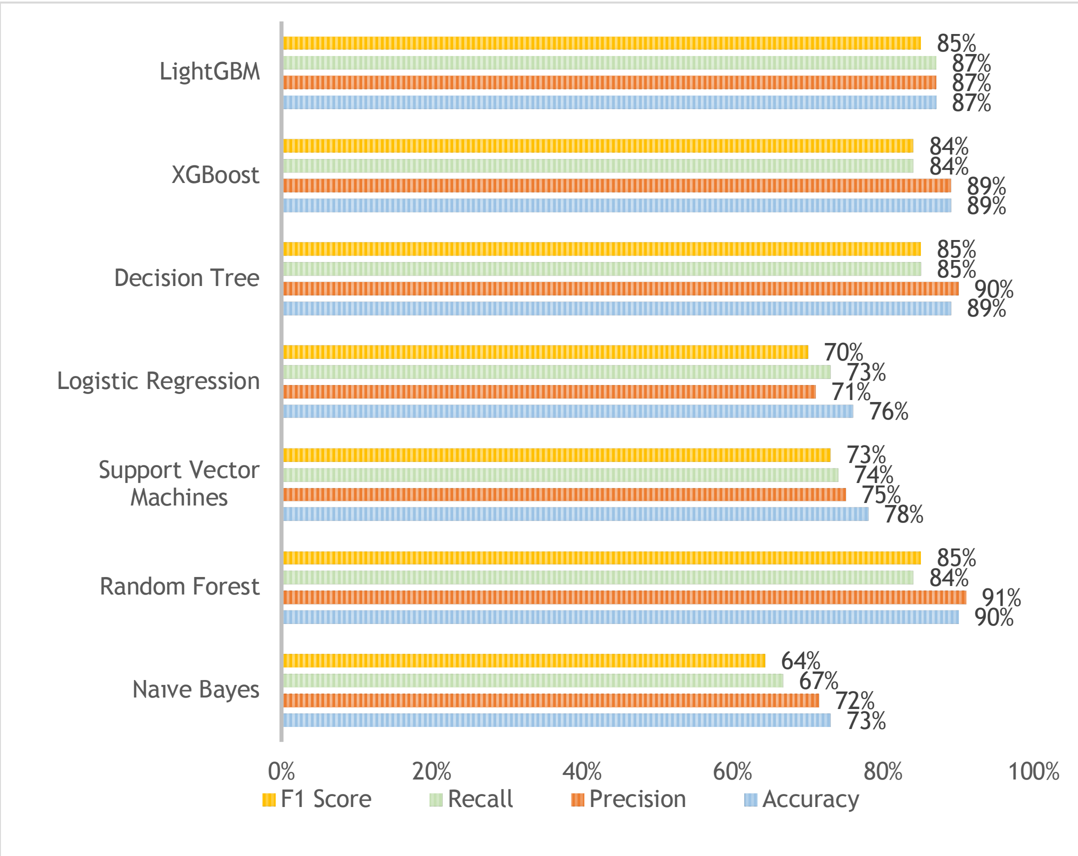

Figure 6. Comparing the performance of algorithms

According to the results given in Figure 6, the Random Forest algorithm gave the best result according to the average of the performance metrics. The accuracy, precision, recall and F1 score values were found to be 0.90, 0.91, 0.84 and 0.85, respectively. The closest successful result to this belongs to the Decision Tree algorithm. The XGBoost and LightGBM were the next most successful algorithms. XGBoost gave better results than LightGBM in terms of accuracy and precision. However, LightGBM had higher recall and f1-score values than the values of XGBoost. The XGBoost algorithm outperform other algorithms for large datasets with sparse set of features [16]. On the other hand, the methods used in LightGBM algorithm provide performance improvement in terms of accuracy, memory consumption and training time, especially for large datasets [17]. Therefore, since our dataset has small number of instances and features, we observed lower performances for these algorithms in our experiments. The most unsuccessful algorithm was found to be Naive Bayes of which performance criteria are 0.73, 0.72, 0.67 and 0.64, respectively. The performance metrics of Random Forest algorithm for each personality feature and their averages are given in Table 5.

Table 5. Results of Random Forest algorithm

Personality	Accuracy	Precision	Recall	F1 Score
Optimistic	0,86	0,78	0,83	0,78
Balanced	0,86	0,93	0,87	0,89
Stressful	0,89	0,76	0,83	0,70
Impatient	0,96	0,94	1,00	0,97
Leery	0,93	0,96	0,95	0,95
Uneasy	0,91	0,89	0,64	0,69
Emotional	0,93	0,97	0,89	0,92
Introvert	0,89	0,94	0,89	0,91
Extrovert	0,88	0,96	0,76	0,83
Calm	0,91	0,92	0,78	0,84
Average	**0,90**	**0,91**	**0,84**	**0,85**

According to Table 5, it is seen that the results of the "Random Forest" algorithm for the "impatient" and " leery " features are slightly more successful than the other features.

5. Conclusion

The dataset was applied to the machine learning algorithms and their performances were compared. Considering the Accuracy, Precision, Recall and F1 Score criteria in Figure 6; NB algorithm was found to be the most unsuccessful algorithm. NB algorithm was followed by SVM and Logistic Regression algorithms, respectively. The RF was found to be the most successful algorithm according to accuracy, precision and f1-score values. On the other hand, LightGBM algorithm had the highest recall score value.

The personality test given by a specialist psychologist was applied to 39 people in total. When the results obtained from these tests were compared with the results obtained from handwriting analysis, it was seen that there was a 73% match. The results of 28 of these tests are the same as those from handwriting analysis. The comparison of our study with other similar studies in the literature in terms of dataset size, extracted features, classification algorithms and accuracy results is given in Table 6.

According to Table 6, most of the studies used different numbers of handwritten datasets that they created. Two of the studies used ready datasets, namely EMNIST and CEDAR. The accuracy rates are ranging from %52 to %100 with varying evaluation criteria based on the extracted features. The features used in the studies are categorized in two classes, namely text layout based and letter shape based. Text layout based features are generally the spaces between words and lines, the margins of the handwritten text. The letter shape based features depend on the letter and its silent characteristic. The other mostly used features were pen pressure and slant of the text.
Machine learning methods applied to the features were mainly ANN, SVM, CNN, k-nearest neighbors (KNN). Apart from these, direct image processing methods have been used to obtain accuracy results in a few studies. There are very few studies using specialist psychologists. In our study, higher accuracy was obtained than most studies in the literature. The accuracy value obtained for RF is 90%. Other values are; 91% precision, 84% recall and 85% f1 score. As a feature study, we will include letterform features in the feature set and increase the size of the dataset to achieve higher accuracy results.

Table 6. Comparison of our study with the literature

Reference	Dataset	Feature Extracted	Methods	Result
Champa Et Al 2010	120 handwriting	Slant of the writing, pen pressure, baseline, letter 't', letter 'y'	Polygonalization method, rule-based technique, gray level threshold value	87.2% average accuracy
Santana Et Al 2010	70 handwriting	Vertical position, pen pressure, spacing between letters, thinning area and letter 'a'	Artificial neural networks and Support Vector Machine	A success rate of 92%
Djamal Et Al 2013	125 handwriting	Spacing between words and lines, zone domination, margin	Multi-structure algorithm and multiple ANN	87-100% using multi-structure algorithm, 52-100% using ANN
Vásquez Et Al 2013	100 handwriting	Slant of the writing, pen pressure, length of initial and final feature, spacing between letters, calligraphic box, and cohesion in writing.	A new method that is digitized and classified by neural networks with a decision fusion block.	A success rate of 94.6%
Fallah And Khotanlou 2016	70 handwritten	Text-dependent features like character size, margin of page, line space, word space	GDA algorithm and MLP neural network	76% accuracy
Sen And Shah 2017	75 handwriting	Word spacing, page margin, baseline, size letter 'i', slant	Image processing techniques	95% accuracy
Gavrilescu And Vizireanu 2018	128 handwriting	Baseline, connecting strokes, word slant, writing pressure, letter 't', letter 'f', Space between the lines	Neural Network	84.4% intra subject, 80.5% inter subject.
Fatimah Et Al 2019	128 handwriting	Space between lines and words, slope, margin, dominant zone, four specific letters ('a', 'g', 's', 't')	Structured approach, CNN	%82.5-100 structured approach, %98 CNN
Chitlangia And Malathi 2019	50 handwriting	Histogram of oriented gradient (HOG)	Support Vector Machine	80% accuracy
Saraswal And.Saxena 2022	EMNIST dataset	Letter size, word spacing, slant of words, pen pressure, baseline, space between words and letter	KNN	89% accuracy
Alamsyah 2022	CEDAR dataset	No feature extraction. A ready dataset is used.	CNN	80.88% accuracy
Our Study	70 handwriting	Page margins, general layout, pen pressure, slant, space between words and lines	Random Forest	90% average accuracy

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

References

[1] E. W. Robertson, *Fundamentals of Document Examination*, Chicago: Nelson-Hall Publishers, 1991.

[2] G. Sheikholeslami, S. N. Srihari, and V. Govindaraju. "Computer aided graphology," Master's thesis, State University of New York, Buffalo, 1995.

[3] H. N. Champa and K. R. AnandaKumar, "Automated Human Behavior Prediction through Handwriting Analysis," *2010 First International Conference on Integrated Intelligent Computing*, Bangalore, India, pp. 160-165, 2010. doi:10.1109/ICIIC.2010.29

[4] S. H. Fatimah, E. C. Djamal, R. Ilyas and F. Renaldi, "Personality Features Identification from Handwriting Using Convolutional Neural Networks," *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia,* pp. 119-124, 2019. doi:10.1109/ICITISEE48480.2019.9003855.

[5] O. Santana, C. M. Travieso, J. B. Alonso and M. A. Ferrer, "Writer identification based on graphology techniques," in *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 6, pp. 35-42, June 2010. doi:10.1109/MAES.2010.5525319

[6] J. L. Vásquez, C. M. Travieso and J. B. Alonso, "Using calligraphies features for off line writer identification," *2013 47th International Carnahan Conference on Security Technology (ICCST), Medellin, Colombia,* pp. 1-6, 2013.

doi:10.1109/CCST.2013.6922062.

[7] A. H. Garoot, M. Safar and C. Y. Suen, "A Comprehensive Survey on Handwriting and Computerized Graphology," *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan,* pp. 621-626, 2017. doi:10.1109/ICDAR.2017.107

[8] A. Anand, D. Patil, S. Bhaawat, S. Karanje and V. Mangalvedhekar, "Automated Career Guidance Using Graphology, Aptitude Test and Personality Test," *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India*, August 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697642

[9] V. Patil and H. Mathur, "A Survey: Machine Learning Approach for Personality Analysis and Writer Identification through Handwriting," *2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India,* pp. 1-5, February 2020. doi:10.1109/ICICT48043.2020.9112449

[10] H. Bacanlı, T. İlhan and S. Aslan, “Beş faktör kuramina dayali bir kişilik ölçeğinin geliştirilmesi: sifatlara dayali kişilik testi (SDKT)”. *The Journal of Turkish Educational Sciences*, vol. 7, no. 2, pp. 261-279, January 2009.

[11] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, 1993.

[12] Z. Wu, W. Lin, Z. Zhang, A. Wen and L. Lin, "An Ensemble Random Forest Algorithm for Insurance Big Data Analysis," *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, *Guangzhou, China*, pp. 531-536, 2017. doi:10.1109/CSE-EUC.2017.99

[13] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958. doi:10.1037/h0042519

[14] B. E. Boser, I. Guyon and V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifiers,” *Proceedings of the 5th Annual Workshop on Computational Learning Theory (COLT’92)*, Pittsburgh, pp. 144-152 July 1992. doi:10.1145/130385.130401

[15] J. Tolles and W. Meurer, “Logistic Regression: Relating Patient Characteristics to Outcomes.,” *JAMA*, vol. 316, no. 5, pp. 533–534. August 2016. doi:10.1001/jama.2016.7653

[16] T. Chen, C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794, August 2016. doi:10.1145/2939672.2939785

[17] K. Guolin, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T. Liu, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 3146–3154, December 2017.

[18] D. Zhang and Y. Gong, "The Comparison of LightGBM and XGBoost Coupling Factor Analysis and Prediagnosis of Acute Liver Failure," in *IEEE Access*, vol. 8, pp. 220990-221003, December 2020. doi:10.1109/ACCESS.2020.3042848.

[19] A. Shehadeh, O. Alshboul, R.E. Al Mamlook, O. Hamedat, "Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression," *Automation in Construction*, vol. 129, pp. 1-16, September 2021. doi:10.1016/j.autcon.2021.103827

[20] E. C. Djamal, R. Darmawati and S. N. Ramdlan, "Application image processing to predict personality based on structure of handwriting and signature," *2013 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), Jakarta, Indonesia,* pp. 163-168, November 2013. doi:10.1109/IC3INA.2013.6819167

[21] B. Fallah and H. Khotanlou, "Identify human personality parameters based on handwriting using neural network," *2016 Artificial Intelligence and Robotics (IRANOPEN)*, *Qazvin, Iran,* pp. 120-126, April 2016. doi:10.1109/RIOS.2016.7529501.

[22] A. Sen and H. Shah, "Automated handwriting analysis system using principles of graphology and image processing," *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, *Coimbatore, India*, pp. 1-6, March 2017. doi:10.1109/ICIIECS.2017.8276061.

[23] M. Gavrilescu and N. Vizireanu, “Predicting the Big Five personality traits from handwriting,” *EURASIP Journal on Image and Video Processing*, vol. 57, July 2018. doi:10.1186/s13640-018-0297-3vol

[24] A. Chitlangia and G. Malathi, “Handwriting Analysis based on Histogram of Oriented Gradient for Predicting Personality traits using SVM,” *Procedia Computer Science*, vol 165, pp. 384-390, 2019. doi:10.1016/j.procs.2020.01.034

[25] A. Saraswal and U. R. Saxena, "Personality Trait Prediction Using Handwriting Recognition with KNN," *2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India,* pp. 551-555, May 2022. doi:10.1109/CISES54857.2022.9844344

[26] D. Alamsyah, Samsuryadi, W. Widhiarso and S. Hasan, "Handwriting Analysis for Personality Trait Features Identification using CNN," *2022 International Conference on Data Science and Its Applications (ICoDSA), Bandung, Indonesia,* pp. 232-238, July 2022. doi:10.1109/ICoDSA55874.2022.9862910

** This article is an extended version of the paper presented at the International Conference on Engineering Technologies (ICENTE'22)*