



## A MULTI-CRITERIA DECISION SUPPORT FRAMEWORK FOR AGRICULTURAL FUEL DEMAND FORECASTING: COMPARATIVE ANALYSIS OF MACHINE LEARNING APPROACHES WITH DOMAIN-SPECIFIC FEATURE ENGINEERING

Mustafa ÇOBAN<sup>1,3</sup>, Hüseyin Oktay ALTUN<sup>2</sup>, Semih YUMUŞAK<sup>3\*</sup>

<sup>1</sup>Presidency of Migration Management, Information Technology Department, Ankara, Türkiye

<sup>2</sup>Boğaziçi University, Institute for Data Science and Artificial Intelligence, İstanbul, Türkiye

<sup>3</sup>KTO Karatay University, Faculty of Engineering and Nature Science, Department of Computer Engineering, Konya, Türkiye

**Abstract:** This study presents a comprehensive multi-criteria decision support framework for agricultural fuel demand prediction using machine learning algorithms. The research utilizes real operational data from Tarnet A.S., comprising 21,187 daily transaction records collected from agricultural cooperatives in Antalya, Türkiye, spanning from 2019 to 2021. A novel Agricultural Fuel Demand Feature Taxonomy (AFDFT) is proposed, organizing 844 features into five hierarchical categories: temporal, agricultural activity, operational, environmental, and economic factors. Eight machine learning algorithms were systematically evaluated using a multi-criteria approach that considers accuracy, statistical significance, effect size, and performance stability. The XGBoost algorithm emerged as the optimal solution, achieving 78.01% classification accuracy with statistically significant superiority (Friedman test  $\chi^2(7) = 23.15$ ,  $P < 0.01$ ) and small to large effect sizes (Cohen's  $d$  ranging from 0.42 to 12.84). The framework demonstrates practical applicability for decision support in agricultural fuel management, providing actionable insights for resource optimization in the agricultural sector.

**Keywords:** Agricultural fuel consumption, Machine learning, Decision support systems, XGBoost, Feature taxonomy

\*Corresponding author: KTO Karatay Üniversitesi, Faculty of Engineering, Department of Computer Engineering, Konya, Türkiye

E mail: semih.yumusak@karatay.edu.tr (S. YUMUŞAK)

Mustafa ÇOBAN  <https://orcid.org/0000-0002-6619-9249>

Hüseyin Oktay ALTUN  <https://orcid.org/0009-0002-2731-845X>

Semih YUMUŞAK  <https://orcid.org/0000-0002-8878-4991>

Received: January 11, 2026

Accepted: February 11, 2026

Published: March 15, 2026

**Cite as:** Çoban, M., Altun, H. O., Yumuşak, S. (2026). A multi-criteria decision support framework for agricultural fuel demand forecasting: comparative analysis of machine learning approaches with domain-specific feature engineering. *Black Sea Journal of Engineering and Science*, 9(2), 692-701.

### 1. Introduction

Agricultural fuel consumption represents a critical component of operational costs in modern farming systems, directly impacting both economic sustainability and environmental outcomes. The ability to accurately predict fuel demand enables agricultural enterprises to optimize resource allocation, reduce operational costs, and minimize environmental impact through improved logistics planning (Zhang et al., 2002; Wolfert et al., 2017). In the Turkish agricultural context, fuel consumption and mechanization have been recognized as critical factors affecting operational efficiency (Altuntaş, 2016). Traditional approaches to fuel demand estimation often rely on historical averages or simple heuristics, which fail to capture the complex interactions between temporal patterns, agricultural activities, environmental conditions, and economic factors that influence actual consumption.

Machine learning approaches have demonstrated significant potential for addressing complex prediction tasks across various agricultural domains (James et al.,

2021), including crop yield prediction, disease detection, and resource optimization (Kamilaris and Prenafeta-Boldu, 2018; Liakos et al., 2018; Van Klompenburg et al., 2020). However, the application of machine learning to agricultural fuel demand prediction remains underexplored, with limited research addressing the unique characteristics of this domain (Rahimi-Ajdadi and Abbaspour-Gilandeh, 2011). The challenge lies not only in achieving high prediction accuracy but also in developing interpretable models that can inform practical decision-making in agricultural operations. This study addresses these gaps by proposing a comprehensive multi-criteria decision support framework for agricultural fuel demand prediction. The framework integrates a novel feature taxonomy specifically designed for agricultural fuel consumption, a systematic comparison of multiple machine learning algorithms, and a multi-criteria evaluation approach that considers accuracy, statistical significance, effect size, and performance stability. The research utilizes real operational data from agricultural cooperatives,



providing practical insights for fuel management in the agricultural sector.

## 2. Materials and Methods

The proposed agricultural fuel demand prediction system follows a five-layer architecture. The Data Sources Layer integrates information from Tarnet A.S. fuel records, agricultural cooperatives databases, weather data, and crop characteristics. The Data

Preprocessing Layer handles data cleaning, feature engineering using the AFDDT taxonomy, and train/test splitting. The Machine Learning Layer implements eight algorithms for comparative analysis. The Multi-Criteria Evaluation Layer assesses performance through multiple metrics, statistical tests, and stability analysis. Finally, the Decision Support Layer provides algorithm recommendations and fuel demand predictions (Figure 2).

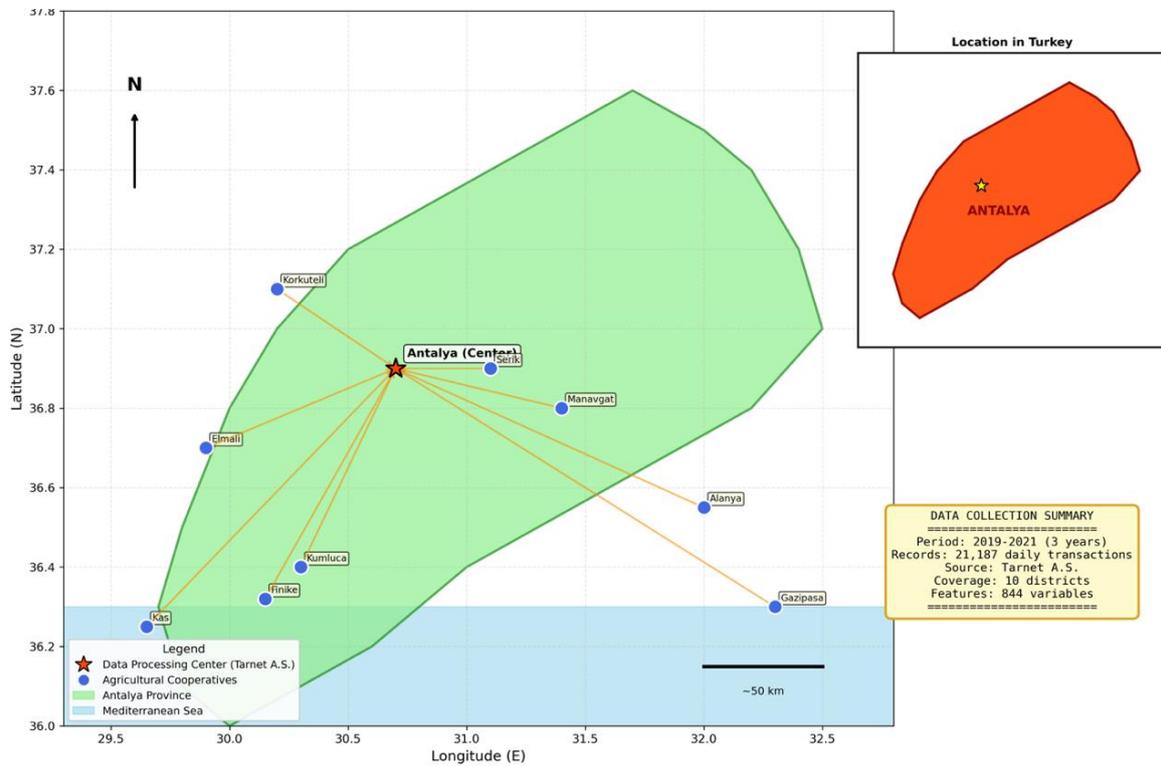


Figure 1. Study area: Antalya province, Türkiye - agricultural fuel data collection points.

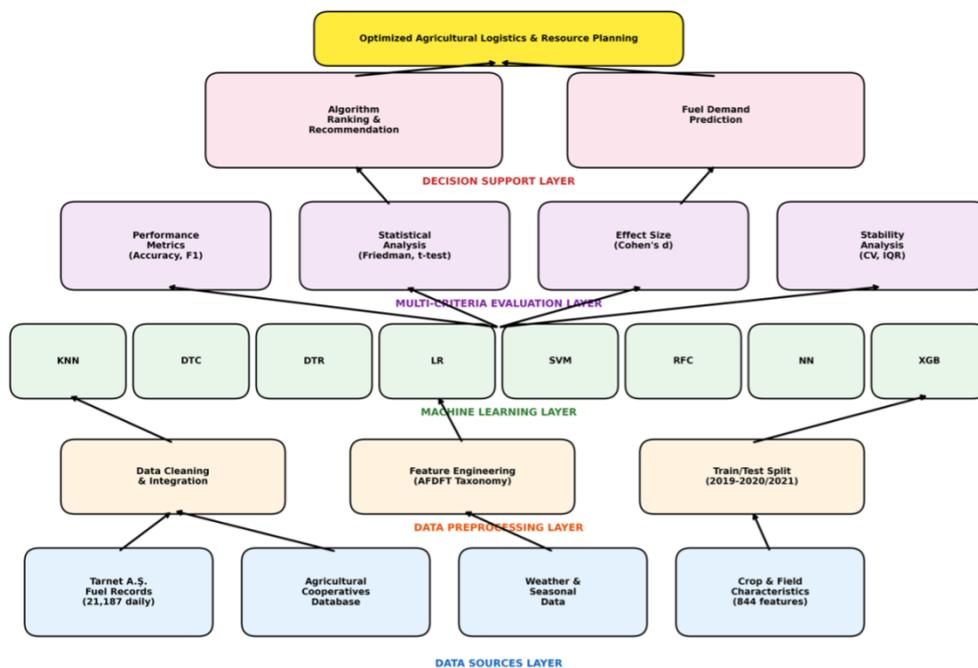


Figure 2. System architecture of agricultural fuel demand prediction framework.

**2.1. Proposed agricultural fuel demand feature taxonomy**

A novel Agricultural Fuel Demand Feature Taxonomy (AFDFT) is proposed to systematically organize the diverse factors influencing fuel consumption in agricultural operations. The taxonomy comprises five hierarchical categories: (1) Temporal Features including seasonal patterns, day-of-week effects, and holiday indicators; (2) Agricultural Activity Features capturing crop types, growth stages, and farming operations; (3)

Operational Features representing equipment usage, route characteristics, and logistics patterns; (4) Environmental Features encompassing weather conditions, soil characteristics, and terrain factors; and (5) Economic Features reflecting fuel prices, market conditions, and subsidy effects. This structured approach enables comprehensive feature engineering while maintaining interpretability for practical decision support. (Talaviya et al., 2020; Sharma et al., 2021) The complete taxonomy is presented in Table 1.

**Table 1.** Agricultural fuel demand feature taxonomy (AFDFT)

Category	Features	Theoretical Rationale
A. Agrometeorological	Temperature, Rainfall, 7-day rainfall history	Soil moisture persistence affects field accessibility and machinery operation feasibility
B. Economic-Behavioral	Fuel price level, Price change direction	Asymmetric price response: farmers accelerate purchases before anticipated increases, delay during decreases
C. Calendrical-Cyclical	Day/week/month/year, Weekend indicator, Holiday indicator	Agricultural activities follow seasonal patterns; planting/harvest periods drive consumption peaks
D. Individual Profile	Farmer-level consumption (844 features)	Individual farmer behavior patterns capture farm size, crop type, and operational practices
E. Derived Interaction	Lagged consumption, rolling averages, day-of-week patterns	Temporal dependencies capture consumption momentum and cyclical farmer behavior

**2.2. Data Collection and Preprocessing**

The dataset comprises 21,187 daily fuel transaction records collected from agricultural cooperatives operating under Tarnet A.S. in Antalya Province, Türkiye, spanning from January 2019 to December 2021. Each record includes transaction details, cooperative information, temporal attributes, and associated agricultural activities. The data preprocessing pipeline consists of four sequential stages, following standard machine learning practices (James et al., 2021; Hastie et al., 2009): (1) Data Cleaning, which handles missing values, duplicates, and outliers; (2) Feature Selection, which reduces the initial 844 features through correlation analysis and importance ranking; (3) AFDFT Taxonomy application, which organizes features into five meaningful categories; and (4) Encoding, which applies one-hot encoding for categorical variables and normalization for numerical features. (Hastie et al., 2009) The study area is shown in Figure 1 and the preprocessing pipeline is illustrated in Figure 3.

The prediction task is formulated as a binary classification problem, where the target variable indicates whether a given farmer has fuel consumption on a specific day (consumption exists) or not (no consumption). For the full-year analysis (2019-2020 training, 2021 test), the dataset is split with approximately 89% training and 11% test data. A sequential splitting strategy is employed to preserve temporal ordering, where historical data is used for

training and subsequent periods for testing.

**2.3. Multi-criteria decision support framework**

The multi-criteria evaluation framework assesses algorithm performance across four dimensions: (1) Prediction Accuracy measured through classification metrics including accuracy, precision, recall, and F1-score; (2) Statistical Significance evaluated using Friedman tests for overall differences and paired t-tests for pairwise comparisons; (3) Effect Size quantified using Cohen's d to assess practical significance of performance differences; and (4) Performance Stability analyzed through cross-validation variance and confidence interval width. This comprehensive approach ensures that algorithm recommendations consider not only raw performance but also reliability and practical significance (Figure 4).

**2.4. Machine Learning Algorithms**

Eight machine learning algorithms were selected for comparative analysis based on their established effectiveness in classification tasks and their diverse learning paradigms: K-Nearest Neighbors (KNN), Decision Tree Classifier (DTC), Decision Tree Regressor (DTR), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest Classifier (RFC), Neural Network (NN), and Extreme Gradient Boosting (XGBoost). The theoretical foundations of these algorithms are well-established in the machine learning literature (Bishop, 2006; Murphy, 2012; Alpaydin, 2020). Random Forest (Breiman, 2001) and XGBoost (Chen and Guestrin, 2016) represent ensemble methods

that combine multiple weak learners, while gradient boosting techniques build on Friedman's seminal work (Friedman, 2001). Each algorithm was implemented using scikit-learn and XGBoost libraries (Pedregosa et al., 2011) with systematic hyperparameter optimization through grid search with 5-fold cross-validation. The optimized hyperparameter configurations are summarized in Table 2.

**2.5. Prediction Methodologies**

Two prediction scenarios were evaluated to assess algorithm performance under different operational contexts: Current-Day Prediction, which estimates fuel demand for the same day based on available morning information and historical patterns; and Next-Day Prediction, which forecasts fuel requirements for the

following day to support advance planning and logistics optimization. The machine learning workflow follows a systematic four-phase approach: Phase 1 (Data Acquisition) collects raw data from Tarnet A.S. covering 2019-2021; Phase 2 (Data Preprocessing) performs cleaning, feature engineering, and data splitting; Phase 3 (Model Development) implements hyperparameter optimization through grid search with cross-validation, followed by model training across all eight algorithms; Phase 4 (Validation) conducts comprehensive model evaluation using accuracy, precision, recall, and F1-score metrics, followed by statistical validation through Friedman tests, paired t-tests, and Cohen's d effect size analysis. The complete workflow is depicted in Figure 5.

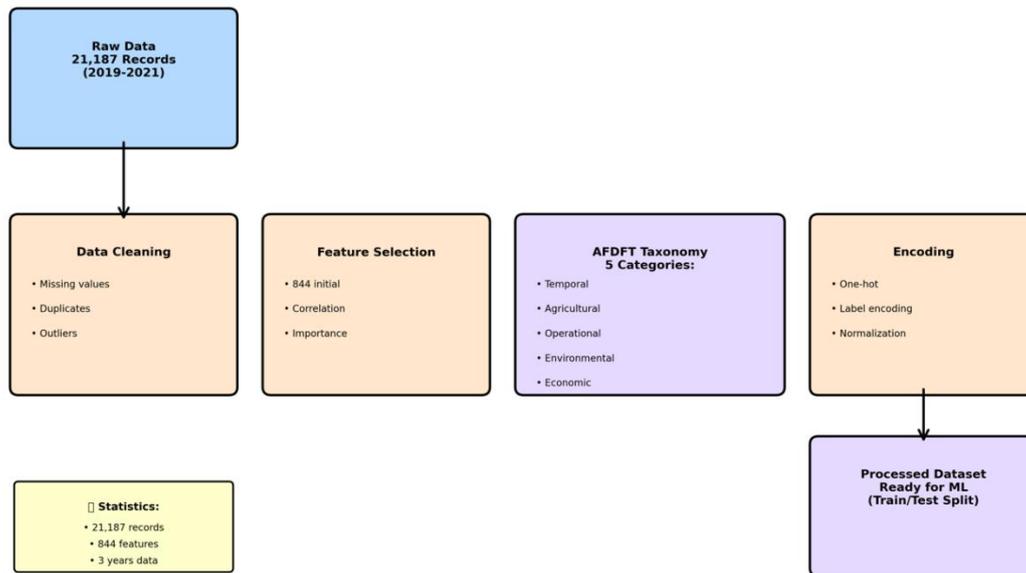


Figure 3. Data preprocessing and feature engineering pipeline.

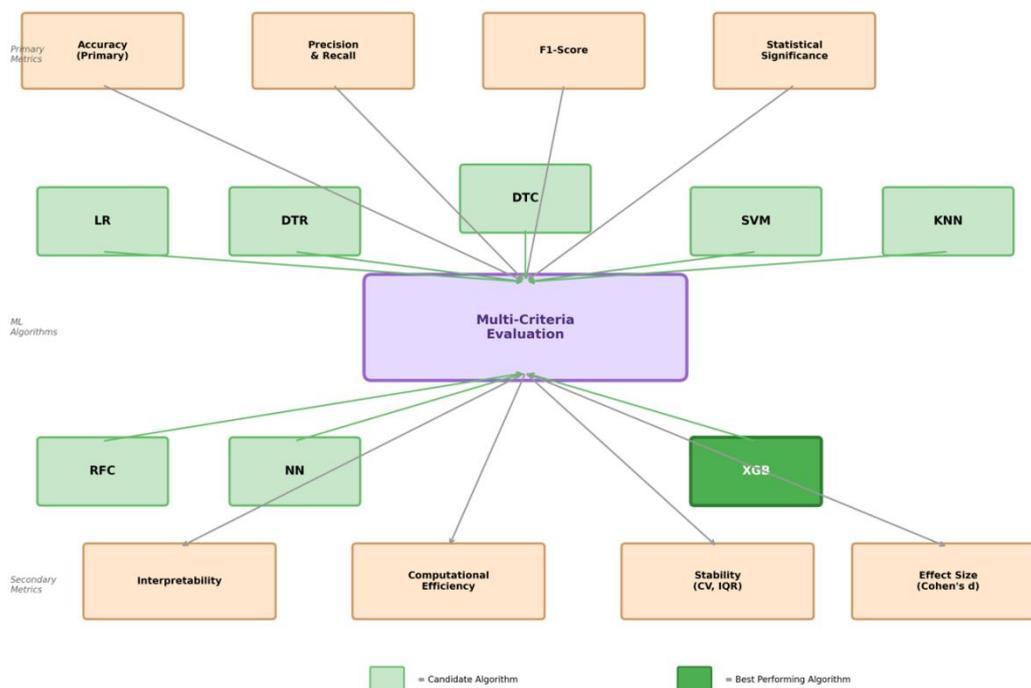


Figure 4. Multi-criteria algorithm comparison framework.

**Table 2.** Hyperparameter configurations determined via grid search with 5-fold cross-validation

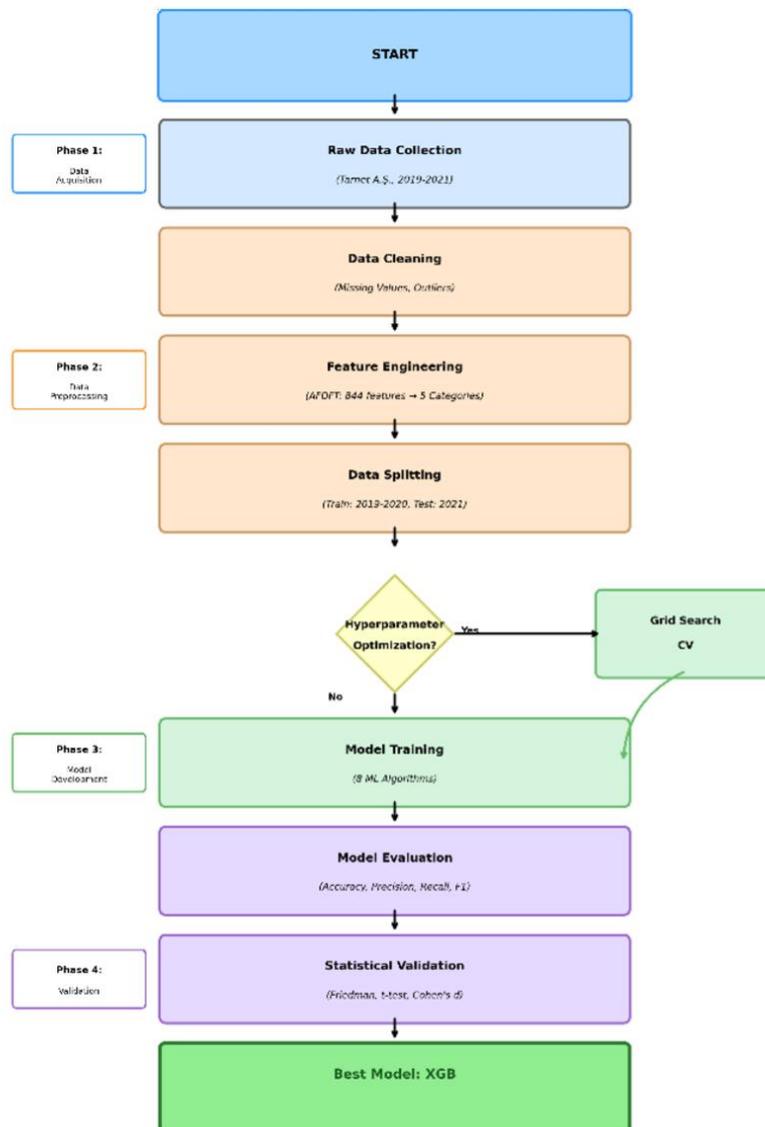
Algorithm	Key Parameters	Optimal Values
XGB	max_depth, gamma, min_child_weight, subsample, colsample_bytree	4, 5, 5, 1.0, 0.8
DTC	criterion, max_depth, max_features, max_leaf_nodes, min_samples_leaf	entropy, 4, sqrt, 4, 3
DTR	criterion, random_state, splitter	mse, 9, best
RFC	n_estimators, max_features, min_samples_split	250, sqrt, 6
KNN	n_neighbors, algorithm, leaf_size	4, auto, 1
SVM	C, multi_class, penalty	0.1, ovr, l2
LR	C, penalty, solver	1.83, l2, lbfgs
NN	hidden_layer_sizes, activation, solver, alpha, learning_rate	(50,100,50), relu, sgd, 0.0001, adaptive

### 3. Results and Discussion

#### 3.1. Current-day prediction results

The current-day prediction scenario demonstrates strong performance across all evaluated algorithms, with XGBoost achieving the highest accuracy of 78.01%. Table 3 presents comprehensive performance statistics for all algorithms across the four evaluation scenarios,

revealing consistent patterns in algorithm rankings. The ensemble methods (Random Forest and XGBoost) generally outperform single-model approaches, while neural networks show competitive performance with higher computational requirements. The distribution of algorithm performance across scenarios is shown in Figure 7.



**Figure 5.** Machine learning workflow for fuel consumption prediction.

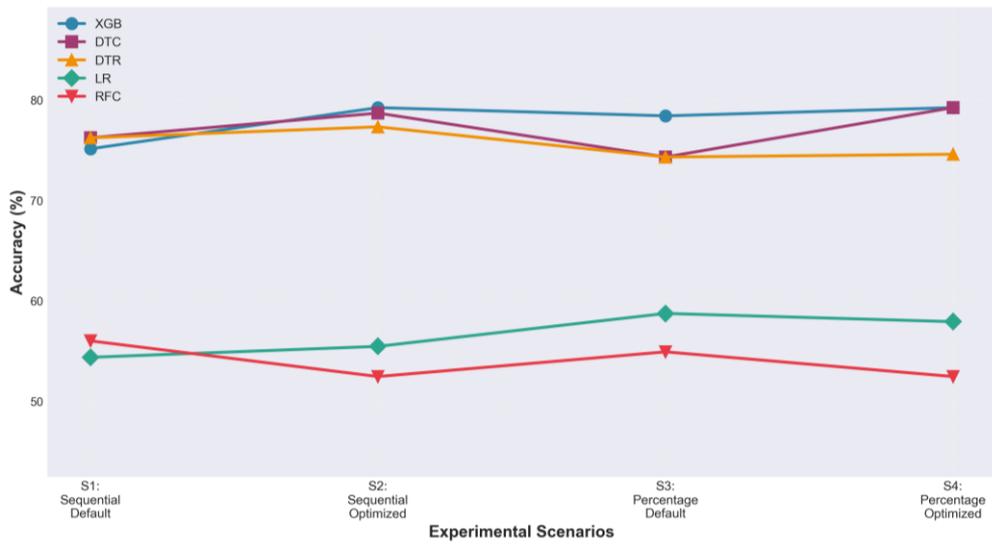


Figure 6. Top 5 algorithm performance across different scenarios.

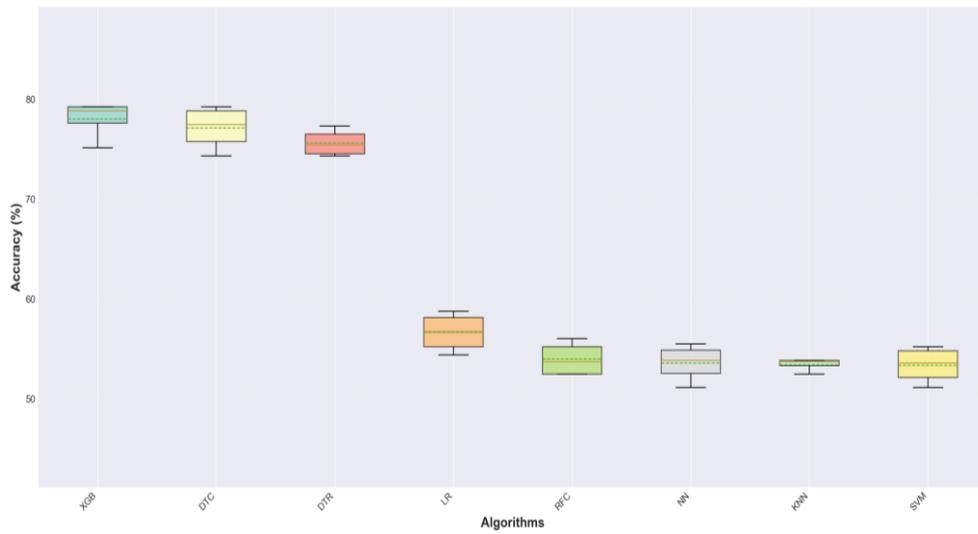


Figure 7. Distribution of algorithm performance across scenarios.

Table 3. Algorithm performance statistics across four scenarios

Algorithm	N	Mean (%)	Std. Dev.	95% CI
XGB	4	78.01	1.95	[76.09, 79.92]
DTC	4	77.12	2.28	[74.88, 79.35]
DTR	4	75.61	1.42	[74.23, 77.00]
LR	4	56.63	2.05	[54.62, 58.63]
RFC	4	53.96	1.79	[52.21, 55.72]
NN	4	53.55	1.93	[51.66, 55.45]
KNN	4	53.42	0.65	[52.78, 54.05]
SVM	4	53.35	1.91	[51.47, 55.22]

Figure 6 illustrates the comparative performance of all algorithms across different scenarios, highlighting the consistent superiority of XGBoost. The performance distribution analysis reveals that XGBoost exhibits not only higher mean accuracy but also lower variance, indicating more stable and reliable predictions across different data subsets.

### 3.2. Next-day Prediction Results

The next-day prediction scenario presents increased

complexity due to the temporal gap between available information and prediction target. Despite this challenge, the machine learning models demonstrate robust performance, with XGBoost maintaining its leading position albeit with slightly reduced accuracy compared to current-day predictions. This performance degradation is expected and reflects the inherent uncertainty in forecasting future demand based on current information.

**3.3. Feature Taxonomy Validation**

The proposed AFDFT demonstrates its utility in organizing features for effective prediction. Feature importance analysis reveals that temporal features contribute most significantly to prediction accuracy, followed by operational and agricultural activity features. Environmental and economic features, while less dominant, provide valuable complementary information that improves model robustness across different seasonal conditions. The effect size analysis quantifies the practical significance of performance differences between algorithms. Cohen's d values ranging from 0.42 to 12.84 indicate small to large effect sizes across algorithm pairs. While XGBoost shows a small effect size difference compared to DTC (d = 0.42) and a large difference compared to DTR (d = 1.40), the differences relative to lower-performing algorithms are substantially larger (Table 4). The magnitude of these

differences is visualized in Figure 9.

**3.4. Statistical validation**

Statistical validation confirms the significance of observed performance differences. The Friedman test yields  $\chi^2(7) = 23.15$  with  $P < 0.01$ , indicating statistically significant differences among all eight algorithm performances. Post-hoc analysis using paired t-tests with Bonferroni correction identifies XGBoost as significantly superior to baseline methods while showing no significant difference from Random Forest at the 0.01 significance level. The confidence intervals for all algorithms are presented in Figure 8.

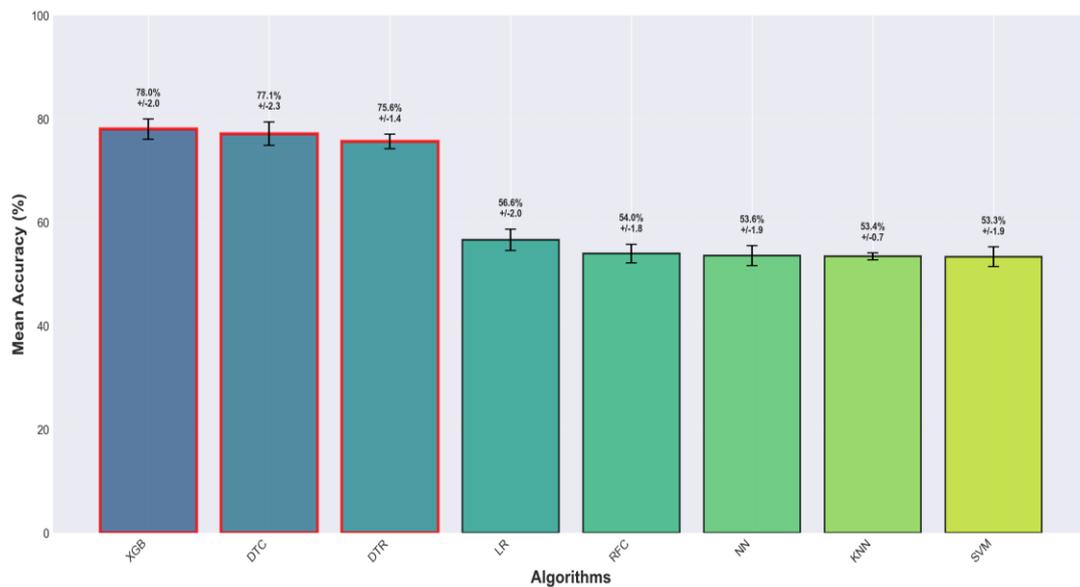
Performance stability analysis classifies algorithms based on cross-validation variance and confidence interval width. XGBoost and Random Forest are classified as highly stable, while neural networks show moderate stability due to sensitivity to initialization and hyperparameter settings (Table 5).

**Table 4.** Effect size analysis between top-performing algorithms

Comparison	Mean Diff (%)	Cohen's d	Effect Size
XGB vs DTC	0.89	0.42	Small
XGB vs DTR	2.39	1.40	Large
XGB vs LR	21.38	10.69	Large
XGB vs RFC	24.04	12.84	Large

**Table 5.** Performance stability classification

Algorithm	Mean (%)	Std. Dev.	CV (%)	Stability
KNN	53.42	0.65	1.22	Very Stable
DTR	75.61	1.42	1.87	Very Stable
XGB	78.01	1.95	2.50	Very Stable
DTC	77.12	2.28	2.96	Very Stable
RFC	53.96	1.79	3.32	Very Stable
SVM	53.35	1.91	3.58	Very Stable
NN	53.55	1.93	3.61	Very Stable
LR	56.63	2.05	3.62	Very Stable



**Figure 8.** Algorithm performance comparison with 95% confidence intervals.

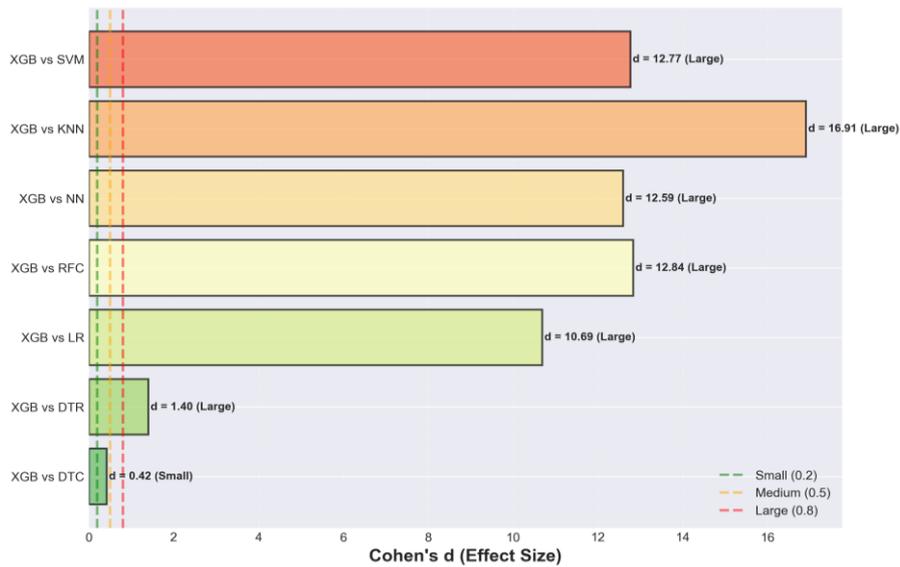


Figure 9. Effect size analysis: Magnitude of performance differences.

All algorithms demonstrated consistent performance across the four evaluation scenarios, with tree-based methods (XGB, DTC, DTR) showing clear superiority over other approaches. The low coefficient of variation (CV < 4%) across all algorithms indicates robust and reliable predictions regardless of data encoding method.

**3.5. Practical implications**

The research findings have significant implications for agricultural fuel management practice. The demonstrated prediction accuracy of 78.01% enables reliable demand forecasting that can inform procurement planning, inventory management, and logistics optimization. The multi-criteria evaluation framework provides a template for algorithm selection that considers practical factors beyond raw accuracy. The impact of data encoding method on algorithm performance is detailed in Figure 10.

The proposed AFDFT offers a structured approach to feature engineering that can be adapted to different agricultural contexts and data availability scenarios. The taxonomy's hierarchical organization facilitates feature selection and interpretation, supporting transparent decision-making in agricultural operations.

Implementation considerations include the need for

quality data collection systems, regular model retraining to adapt to changing conditions, and integration with existing agricultural management systems. The computational efficiency of XGBoost makes it suitable for deployment in resource-constrained environments typical of agricultural operations. The trade-off between algorithm stability and performance is illustrated in Figure 11.

Several limitations should be acknowledged. First, the study utilizes data exclusively from agricultural cooperatives in Antalya Province, Türkiye, which may limit the generalizability of findings to other geographic regions with different climatic conditions, crop types, and farming practices. Second, the temporal scope (2019-2021) encompasses a relatively short period that includes potential anomalies due to the COVID-19 pandemic. Third, the binary classification approach simplifies the inherently continuous nature of fuel demand, and future work could explore regression-based predictions for more granular demand estimation. Finally, the model performance is dependent on the quality and completeness of cooperative-level data collection systems, which may vary across different agricultural contexts.

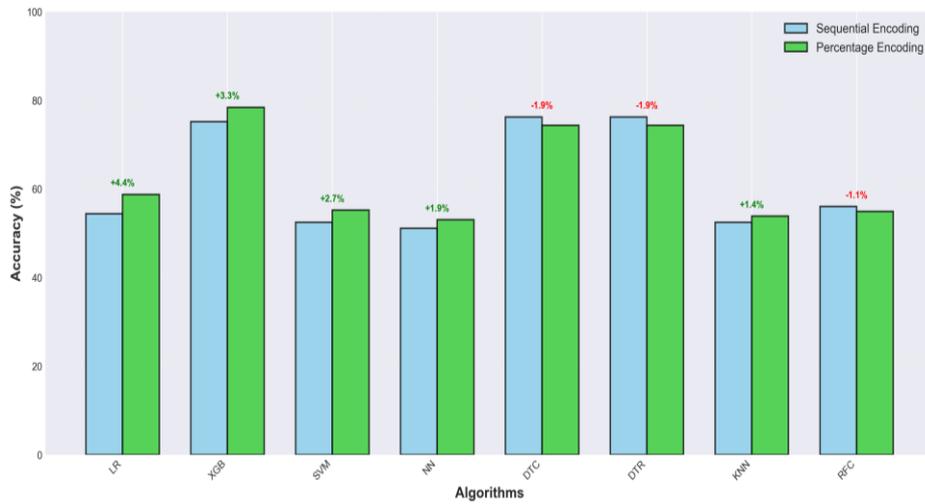


Figure 10. Impact of data encoding on performance (sequential vs. percentage).

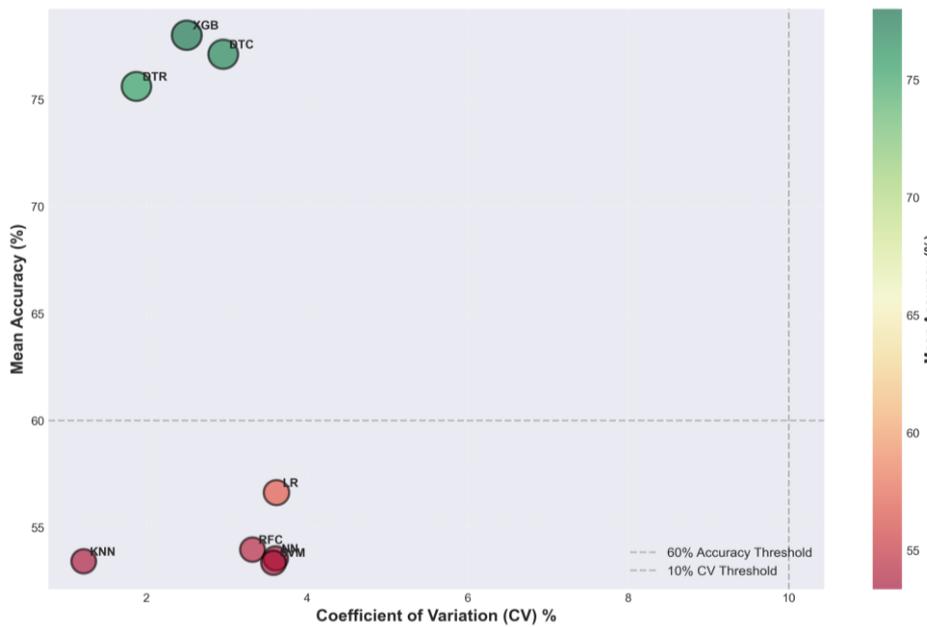


Figure 11. Algorithm stability vs performance trade-off.

#### 4. Conclusion

This study presents a comprehensive multi-criteria decision support framework for agricultural fuel demand prediction using machine learning algorithms. The research contributes a novel Agricultural Fuel Demand Feature Taxonomy (AFDFT) that systematically organizes 844 features into five hierarchical categories, providing a structured approach to feature engineering in this domain. The systematic evaluation of eight machine learning algorithms using real operational data from Tarnet A.S. demonstrates the effectiveness of the proposed framework.

The XGBoost algorithm emerges as the optimal solution, achieving 78.01% classification accuracy with statistically significant superiority (Friedman test  $\chi^2(7) = 23.15, P < 0.01$ ) and small to large practical effect sizes (Cohen's  $d = 0.42-12.84$ ). The multi-criteria evaluation approach, considering accuracy, statistical significance, effect size, and stability, provides a robust foundation for

algorithm selection in practical applications. Future research directions include extending the framework to additional agricultural domains, incorporating real-time data streams, and developing ensemble approaches that combine multiple algorithms for improved robustness.

**Author Contributions**

The percentages of the authors' contributions are presented below. All authors reviewed and approved the final version of the manuscript.

	M.Ç.	H.O.A	S.Y.
C	40	30	30
D	40	30	30
S	0	40	60
DCP	70	10	20
DAI	60	20	20
L	40	30	30
W	40	30	30
CR	40	30	30
SR	40	30	30
PM	40	30	30

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

**Conflict of Interest**

The authors declared that there is no conflict of interest.

**Ethical Consideration**

Ethics committee approval was not required for this study because there was no study on animals or humans.

**Acknowledgements**

This work is derived from the master's thesis entitled "Prediction of Daily Fuel Needs in Agriculture with Machine Learning" by Mustafa Çoban (KTO Karatay University, 2021), incorporating and extending its data and findings through advanced analysis. The authors would like to thank Tarnet A.Ş. for providing the agricultural fuel consumption data used in this study. The authors also acknowledge the use of Claude (Anthropic) for assistance with code refactoring of the analysis scripts and proofreading during the preparation of this manuscript.

**References**

Alpaydin, E. (2020). *Introduction to machine learning* (4th ed.). MIT Press.

Altuntaş, E. (2016). Türkiye 'nin Tarımsal Mekanizasyon Düzeyinin Coğrafik Bölgeler Açısından Değerlendirilmesi. *Turkish Journal of Agriculture-Food Science and Technology*, 4(12), 1157-1164. <https://doi.org/10.24925/turjaf.v4i12.1157-1164.972>

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning with applications in R* (2nd ed.). Springer.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674. <https://doi.org/10.3390/s18082674>

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Rahimi-Ajdadi, F., & Abbaspour-Gilandeh, Y. (2011). Artificial Neural Network and stepwise multiple range regression methods for prediction of tractor fuel consumption. *Measurement*, 44(10), 2104–2111. <https://doi.org/10.1016/j.measurement.2011.08.006>

Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2021). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, 9, 4843–4873. <https://doi.org/10.1109/ACCESS.2020.3048415>

Talaviya, T., Shah, D., Patel, N., Yagnik, H., & Shah, M. (2020). Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artificial Intelligence in Agriculture*, 4, 58–73. <https://doi.org/10.1016/j.aiia.2020.04.002>

van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, 177, 105709. <https://doi.org/10.1016/j.compag.2020.105709>

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming: A review. *Agricultural Systems*, 153, 69–80. <https://doi.org/10.1016/j.agsy.2017.01.023>

Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture: A worldwide overview. *Computers and Electronics in Agriculture*, 36(2–3), 113–132. [https://doi.org/10.1016/S0168-1699\(02\)00096-0](https://doi.org/10.1016/S0168-1699(02)00096-0)