

## PREDICTION USING THE SVM ALGORITHM IN MACHINE LEARNING

**Qarshiyev Abduvali Berkinovich**

*Candidate of Physical and Mathematical Sciences, Professor Professor at the Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi ORCID: 0000-0001-6121-9928*

**Tursunov Muhammadsolih Sa'din o'g'li**

*Doctor of Philosophy (PhD) in Philological Sciences, Associate Professor Associate Professor at the Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi E-mail: muhammadsolih927@gmail.com ORCID: 0000-0002-6485-3630*

**Quvondiqov Muhammad Dilmurod o'g'li**

*Student at the Samarkand branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi E-mail: quvondiqovmuhammad07@gmail.com ORCID: 0009-0008-1074-3437*

**Abstract:** *This article explores the process of developing a system that predicts users' purchasing probability based on their socio-economic indicators using the Support Vector Machines (SVM) algorithm. The study covers the issues of data training, scaling, and integrating the trained model into a web interface via the Flask framework.*

**Keywords:** *Machine learning, SVM algorithm, prediction, classification, data analysis, Flask, web integration.*

**Аннотация:** *В данной статье исследуется процесс разработки системы, прогнозирующей вероятность покупки пользователями на основе их социально-экономических показателей с использованием алгоритма метода опорных векторов (SVM). В ходе исследования освещены вопросы обучения данных, масштабирования и интеграции готовой модели в веб-интерфейс через фреймворк Flask.*

**Ключевые слова:** *Машинное обучение, алгоритм SVM, прогнозирование, классификация, анализ данных, Flask, веб-интеграция*

Today, predictive analytics of customer behavior is one of the most important issues in the fields of marketing and e-commerce. In the context of increasing data volumes, artificial intelligence algorithms, particularly classification models, are showing high efficiency in identifying potential customers [3].

Support Vector Machine is a powerful supervised learning algorithm used to classify linear and non-linear data. The main goal of the model is to find the most optimal dividing line, i.e., a hyperplane, among points belonging to different classes (e.g., "Buyers" and "Non-buyers") [1].

From a mathematical point of view, the hyperplane is represented by the following linear equation:

$$\omega \times X + b = 0$$

Here:  $\omega$  – is the weight vector perpendicular to the hyperplane,  $X$  – is the input parameter vector (in our case, age and salary), and  $b$  is the bias indicator.

To maximize the distance (margin) between the two classes, SVM solves the following optimization problem:

$$\min \frac{1}{2} \|\omega\|^2$$

Here, the model finds such values of  $\omega$  and  $b$  that the gap between the support vectors (the data points closest to the boundary) is as wide as possible. This feature allows the SVM algorithm to make more accurate decisions on new, previously unseen data.

Dataset and its analysis. To test the theoretical model in practice, a dataset prepared based on advertising data was used. This dataset aims to predict the target variable by analyzing the age and salary of customers.

Sample data table used in the study:

Age	Salary (UZS/unit)	Purchased (1-Yes, 0-No)
19	19,000	0
35	20,000	0
26	43,000	0
32	150,000	1
47	25,000	1
48	29,000	1

Model training and visualization. The model was developed using the scikit-learn library of the Python programming language [2]. In order to reduce the variance between the data and ensure the more accurate operation of the algorithm, a standardization process was carried out using StandardScaler. A linear kernel was used during the training process.

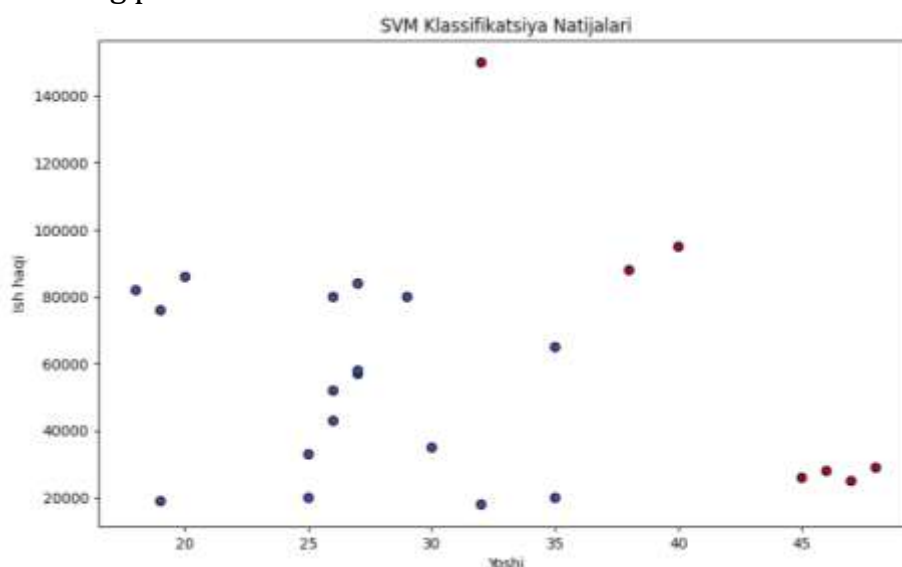


Figure 1. Data classification and the separating hyperplane using the SVM algorithm.

As can be seen in the figure, the model successfully separated the data points into two groups based on age and income. Red dots represent customers who did not buy the product, and blue dots represent potential buyers.

Analysis of results. To test the trained model in real-time, a web application was created based on the Flask framework [4]. The system receives age and salary data from the user and immediately outputs the prediction result using the saved .pkl model file. For example, when a user entered an age of 40 and a salary of 55,000, the system provided the conclusion "Will buy". This process allows marketing analysts to draw quick conclusions.

In conclusion, the conducted research shows that the SVM algorithm has high classification accuracy when working with socio-economic data. The developed algorithm and the method of its integration into a web system can provide practical assistance to enterprises in segmenting their customer base [5]. In the future, it is planned to compare this approach with Deep Learning networks.

#### **REFERENCES:**

1. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
2. Pedregosa, F., et al. "Scikit-learn: Machine learning in Python." *Journal of machine learning research*, 12, 2825-2830 (2011).
3. Provost, F., & Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.
4. Grinberg, M. (2018). *Flask Web Development: Developing Web Applications with Python*. O'Reilly Media.
5. Collection of scientific articles on information technology of the Samarkand branch of TUIT named after al-Khwarizmi (2025).