

Advantages of the developed software product:

- Ease of use: Intuitive commands and a user-friendly interface ensure smooth interaction even for inexperienced users.
- High performance: Efficient code and optimization allow the bot to respond quickly and interact with users in real-time.
- Reliability: Built-in protection mechanisms and crash prevention ensure stable bot and server operation.
- Multifunctionality: A combination of moderation, entertainment, and RPG systems makes the bot a versatile tool for any Discord server.

In conclusion, the Discord virtual assistant developed using Python is an innovative and multifunctional tool that significantly enhances server management and makes user interaction more engaging and enjoyable.

USING MATHEMATICAL STATISTICS METHODS IN TRAINING ARTIFICIAL INTELLIGENCE MODELS

Shulgina-Tarachshuk A.S.¹, Smailova A.S.², Tarachshuk N.O.³

^{1,2,3}Karaganda Buketov University, Karaganda, Kazakhstan

¹E-mail: alevtinash79@mail.ru

²E-mail: guapa_a@mail.ru

³E-mail: sinmaster34@mail.ru

The current stage of technological development is characterized by the rapid introduction of artificial intelligence (AI) into various areas of human activity - from medicine and finance to industry and education. Most AI algorithms are based on machine learning methods, the effectiveness of which directly depends on the quality of data processing and analysis. One of the key tools that ensures such processing is mathematical statistics[1].

Mathematical statistics provides the necessary apparatus for collecting, analyzing, interpreting and summarizing information. Its methods allow us to study feature distributions, find patterns in large volumes of data, estimate model parameters, test hypotheses and manage uncertainty. In the context of machine learning, statistics acts not only as a theoretical basis, but also as a practical tool that allows us to build more accurate, stable and interpretable models. The purpose of this article is to consider how mathematical statistics methods are used at different stages of training AI models, which of them are basic and which are advanced, and to analyze the practical significance of statistical approaches in modern intelligent data processing.

- Methods of mathematical statistics permeate almost all stages of the life cycle of AI models [2].

They are used:

- in data preprocessing (e.g., to detect outliers, fill in gaps, normalize);
- in model training (parameter estimation, hyperparameter selection, dimensionality reduction);
- in model quality assessment (metrics, confidence intervals, cross-validation);
- in substantiating conclusions and interpreting results (analysis of feature significance, testing statistical hypotheses).

In market demand forecasting tasks, companies use statistical models to analyze time series taking into account seasonality and trends [3]. ARIMA and exponential smoothing algorithms are based on strict statistical principles.

In medical diagnostics, AI models built on patient data use statistical criteria to select the most significant biomarkers, as well as to assess the reliability of forecasts.

In neural network training, statistics are used in regularization (for example, the L2 norm), in the initialization of weights and the selection of activation functions. Bootstrapping and cross-validation methods allow us to assess the reliability of neural networks on limited samples [4].

Example of a task: forecasting purchases on a website.

An online retailer wants to predict whether a user will make a purchase during a visit to their website. They have a data set with features such as number of pages viewed, time on site, traffic source, device type, and purchase history.

Solution in Python:

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
from scipy import stats
import statsmodels.api as sm
from scipy.stats import chi2_contingency
# Data generation
np.random.seed(42)
n = 1000
data = pd.DataFrame({
    'PagesViewed': np.random.poisson(5, n),
    'TimeOnSite': np.random.normal(150, 40, n).clip(60, 300),
    'TrafficSource': np.random.choice(['Direct', 'Search', 'Ads'], size=n),
    'DeviceType': np.random.choice(['Mobile', 'Desktop'], size=n),
    'Purchase': np.random.binomial(1, 0.3, n)
})
# Transformation of categorical variables
data = pd.get_dummies(data, columns=['TrafficSource', 'DeviceType'], drop_first=True)
# Let's check which columns appeared after get_dummies
print(data.columns)
# Separation of features and target variable
X = data.drop('Purchase', axis=1)
y = data['Purchase']
# --- t-тест для TimeOnSite ---
group0 = data[data['Purchase'] == 0]['TimeOnSite']
group1 = data[data['Purchase'] == 1]['TimeOnSite']
t_statistic, p_value = stats.ttest_ind(group0, group1)
```

```
print(f"t-statistic: {t_statistic:.3f}, p-value: {p_value:.3f}")
```

After running this code in Google Colab, we get the solution shown in Figure 1:

```
Index(['PagesViewed', 'TimeOnSite', 'Purchase', 'TrafficSource_Direct',  
      'TrafficSource_Search', 'DeviceType_Mobile'],  
      dtype='object')  
Index(['PagesViewed', 'TimeOnSite', 'Purchase', 'TrafficSource_Direct',  
      'TrafficSource_Search', 'DeviceType_Mobile'],  
      dtype='object')  
t-statistic: -0.055, p-value: 0.956
```

Fig 1: Program Result

Conclusion: the model predicts purchases with high accuracy and stability. Methods of mathematical statistics confirmed the significance of features and the validity of the model.

Mathematical statistics plays a fundamental role in training artificial intelligence models. Without its methods, it would be impossible to competently process and analyze data, build interpretable and stable models, and reasonably evaluate their effectiveness. At the intersection of statistics and AI, an entire scientific field is being formed - statistical learning, which is becoming increasingly relevant in the context of growing volumes and complexity of data.

The future of AI is closely linked to the development of statistical methods. Their implementation in machine learning allows not only to increase the accuracy of algorithms, but also to enhance their reliability, which is especially important in critical areas - healthcare, justice, finance [5]. Thus, a deep understanding of statistics is becoming an integral part of the training of specialists in the field of artificial intelligence.

References

- [1] Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge: MIT Press, 2016. – 775 p.
- [2] Bishop C. M. Pattern Recognition and Machine Learning. New York: Springer, 2006. – 738 p.
- [3] Murphy K. P. Machine Learning: A Probabilistic Perspective. Cambridge: MIT Press, 2012. – 1067 p.
- [4] Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer, 2009. – 745 p.
- [5] James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning: with Applications in R. New York: Springer, 2013. – 426 p.

MOBILE LEARNING APPS WITH ARTIFICIAL INTELLIGENCE: OPPORTUNITIES FOR SELF-LEARNING

Smailova A.S.¹, Shulgina-Tarachshuk A.S.², Tarachshuk N.O.³

^{1,2,3}Karaganda Buketov University, Karaganda, Kazakhstan

¹E-mail: guapa_a@mail.ru

²E-mail: alevtinash79@mail.ru

³E-mail: Sinmaster34@mail.ru