

- A high VC-dimension does not always lead to overfitting if the data has a specific structure (e.g., low effective dimensionality).

Model theory offers a powerful formal language for analyzing classes of machine learning models. VC-dimension and the NIP property serve as key indicators of a model's generalization ability and can be used for theoretically grounded algorithm selection based on data volume and interpretability requirements.

The further development of interdisciplinary approaches – at the intersection of logic, probability theory, and machine learning – may lead to the creation of new, explainable, and provably robust ML systems.

References

- [1] Kaplan, I., Ramsey, N. (2023). Model Theory and Machine Learning: A Survey. <https://arxiv.org/abs/2301.08731>
- [2] Vapnik, V. N., Chervonenkis, A. Y. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications*, 16(2), 264–280.
- [3] Kaplan, I., Shelah, S. (2020). On the structure of NIP theories. *Journal of Symbolic Logic*, 85(4), 1623–1650. <https://doi.org/10.1017/jsl.2020.61>
- [4] Livni, R., Simon, P. (2020). VC Theory in Machine Learning: A Model-Theoretic Perspective. arXiv:2005.04372. <https://arxiv.org/abs/2005.04372>
- [5] Guide to NIP Theories. Retrieved from <https://math.univ-lyon1.fr/~simon/guideNIP.pdf>

DEVELOPMENT OF A VIRTUAL ASSISTANT USING PYTHON

Samoilova Irina Alexeevna¹, Spirina Yelena Aleksandrovna², Smirnova Marina Aleksandrovna³, Popova Nadegda Viktorovna⁴

^{1,2,3,4}Karaganda Buketov University, Karagandy, Kazakhstan

¹E-mail: irinasam2005@mail.ru

²E-mail: sea_spirina@mail.ru

³E-mail: smirnova_marina_alex@mail.ru

⁴E-mail: dandn@mail.ru

This article aims to explore and develop technologies for creating a virtual assistant using the Python programming language, with the Discord online platform as a case study.

Python was chosen as the development tool because it is a powerful and versatile programming language widely used in various areas of software development, including the creation of bots (virtual assistants, conversational agents) for online platforms. The objective of this work was to develop an innovative solution that meets the needs of modern users and contributes to the development of online communities on the Discord platform.

Virtual assistants on Discord are programs designed to interact with users in an online environment. These bots possess various functions and capabilities that allow them to participate in discussions, provide information, entertain, and assist server members. Bots are software applications created to automate tasks and perform specific actions within chats. The role of a

virtual assistant in Discord may vary depending on users' goals and needs. This research examines the role of virtual assistants on Discord and their impact on communication processes.

When selecting technologies and tools for developing a virtual assistant for Discord using Python, several key factors must be considered:

Libraries for working with the Discord API: A primary criterion is the availability of libraries that allow interaction with the Discord API in Python. `discord.py` is one of the most popular options, providing a convenient and efficient interface.

Alternative libraries: In addition to `discord.py`, there are other Python libraries for interacting with the Discord API, such as `aiogram`. Their capabilities must be analyzed to choose the most appropriate option depending on the project's requirements.

Frameworks and development tools: To build a bot effectively in Python, various frameworks and tools can be used, such as Flask or FastAPI for web servers, and PyCharm or VS Code as integrated development environments.

Development and debugging tools: High-quality development requires tools that offer rich functionality and ease of use when working with code.

Version control system: Using a version control system like Git is a crucial aspect of any software project. It allows tracking changes, collaborating with other developers, and restoring previous versions if needed.

Additional libraries and tools: Depending on the bot's required features, additional libraries may be necessary, such as those for working with databases or handling asynchronous requests.

Currently, there are many bots on Discord with diverse functionalities, developed in various programming languages. However, Python remains one of the most popular choices for bot development due to its simplicity and powerful libraries like `discord.py`.

Natural Language Processing (NLP) plays a key role in developing virtual assistants for Discord. NLP algorithms allow bots to understand and analyze user messages, enabling them to perform a range of functions, from answering questions to managing server tasks.

Developing NLP algorithms requires a deep understanding of text data and expertise in machine learning and language processing. Python offers extensive support in this area through libraries and frameworks such as NLTK, spaCy, and TensorFlow.

Successful development of a Discord virtual assistant (bot) in Python involves following specific stages, including creating a prototype. Prototyping helps evaluate the bot's functionality and receive user feedback.

The development process consists of several essential stages, each important for building a fully functional and reliable bot. This thesis discusses the entire process from concept to testing and deployment. The key stages are:

1. Defining the task and requirements.
2. Selecting technologies and tools.
3. Designing the bot's architecture.
4. Implementing core features.
5. Testing and debugging.
6. Deployment and configuration.
7. Maintenance and updates.

An important part of development is post-deployment support and regular updates. This includes monitoring bot performance, fixing bugs, adding new features, and adapting to Discord API changes. Regular updates help maintain the bot's relevance and stability.

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).