



# Kseniia Vaniushkina

Machine Learning Engineer

[GitHub](#)

@justavk

+82 10 6514 2124

averksuu@naver.com

## ABOUT ME

Machine Learning Engineer experienced in developing models for analyzing medical, textual, and audio data.

Worked with XGBoost, LightGBM, neural networks, and RAG-based approaches, including model integration into Telegram bots.

Proficient in Python, TensorFlow, FastAPI, LangChain, and tools for tabular, time-series, and visual data.

Quick learner focused on practical, high-impact results.

## SKILLS

Programming Languages: Python, SQL

Data Analysis: NumPy, Pandas, Matplotlib, scikit-learn, tqdm, pickle, AutoML, AutoKeras, Auto-TS, TimeseriesGenerator

CV: OpenCV, PIL/Pillow, SciPy, MoviePy, VidGear, TerraSegmentation, TerraYolo (Ultralytics), DeepSort, EasyOCR, diffusers

NLP/ audio: GPT (OpenAI), HuggingFace, T5, BERT, LangChain, Faiss, SpeechRecognition, LibROSA, PyAudio, Jiwer, Google Text-to-Speech, Yandex SpeechKit

Infrastructure & Deployment: Docker, FastAPI, os, time, gdow

Google Colab, Git, google cloud

## EDUCATION

2025

University of Artificial Intelligence  
Data Science, Neural Networks, Machine Learning & Artificial Intelligence

Coursera I DeepLearning.AI & Stanford University

- Machine Learning Specialization
- Natural Language Processing Specialization

Yandex ML Bootcamp  
«Training 3.0»

## LANGUAGES

English - C1

Korean - C1

## WORK EXPERIENCE

04.2025 - 07.2025

AI System for Analyzing and Automating Client Feedback and Requests on Marketplaces (with operator performance evaluation)

LLC "Hyper"

- Built a multi-output XGBoost model to predict ADHD diagnosis and gender using fMRI data and questionnaires.
- Applied SMOTE to balance class distributions and improve classification quality.

01 - 05. 2025

Project «Unraveling the Mysteries of the Female Brain: Sex Patterns in ADHD»

Kaggle WiDS Datathon

- Built a multi-output XGBoost model to predict ADHD diagnosis and gender using fMRI data and questionnaires.
- Applied SMOTE to balance class distributions and improve classification quality.

Tech Stack: Python, Pandas, NumPy, scikit-learn, XGBoost, SMOTE, Matplotlib, Seaborn

Link

11.2024 - 01.2025

AI System for Analyzing Continuous Glucose Monitoring (CGM) Data to Assess Diabetes Treatment Effectiveness

LLC "Endocrinology Online"

- Analyzed CGM, insulin therapy, diet, and physical activity data.
- Built LightGBM and XGBoost models using sliding window techniques.
- Developed a Telegram bot for real-time predictions, user data storage, and on-demand model fine-tuning.

Tech Stack: Python, Pandas, NumPy, scikit-learn, LightGBM, XGBoost, Telegram Bot API, python-telegram-bot, pickle, sliding window approach

Link

01 - 02.2025

Pet Project: "Otitis Image Classifier"

- Built a binary CNN-based classifier for otoscopic image analysis.
- Preprocessed input data (128×128 px), trained and validated the model.
- Achieved 75% accuracy on the test set.

Tech Stack: Python, TensorFlow, Keras, CNN, NumPy, Matplotlib, PIL

Link

2021 - 2024

Korean & English Language Instructor / Translator

- Conducted over 3,000 individual and group lessons.
- Trained more than 100 students for the TOPIK exam and university admissions in Korea.