Irina Nosal

 Irina-Na | irinanosal | ironia.i@gmail.com | +381 65 2655159 | +7(922)279-88-56

SUMMARY

ML engineer and Data Scientist with Ph.D., expertise in NLP, AI, matching, search, and LLM for business decisions. 14+ years in Tech (Telecom, InfoSec, ML), 8+ years in software dev (Python, SQL, Matlab, C, Assembler), 3+ years commercial. 5+ years in research, 3+ years in Data Science, 1+ year in ML product management. Has 8+ publications. Hirsh Index 4.

WORK EXPERIENCE

SENIOR ML ENGINEER — HAQQ Network   Mar 2024 - present

Blockchain ecosystem HAQQ with capital of $200 million and more then 200 employees
– Developed system design on API microservices, intent detection, and generative components for RAG.
– Management the gathering requirements for AI ecosystem, coordinating the work of several teams.
– Designing quality evaluation tools for Retrieval-Augmented Generation (RAG), lead ML QA work.
Tools: LLMs, OpenAI, LangChain, HuggingFace, FastAPI, PostgreSQL, Gensim, Spacy, RAGAS.

DATA SCIENTIST, ML ENGINEER — AUTODOC EU   Oct 2022 - Mar 2024

EU top **car-parts e-store** with €1.1B in sales and 5.2M products.
Operating in 27 EU countries, 5000+ employees at 8 sites
– Develop and evaluate machine learning models, work closely with data engineers to extract, transform, and load (ETL) data.
– Conduct analytical research, ad-hoc analytics, and build dashboards; perform data preprocessing, exploratory data analysis (EDA), and hypothesis testing; implement and test solutions.
– Collaborate with cross-functional teams to translate business requirements into technical solutions.
– Process and analyze textual data from various sources, Design and plan system architecture.
Tools: Azur OpenAI API, LangChain, Glove, DBSCAN, PCA/TSNE, HuggingFace, GCP, Airflow, SQL, Metabase.

ML ENGINEER — WELLNITE   Dec 2021 - Oct 2022

Well-being **self-care** and **medtech** startup hit $17.6M revenue in 2023 with a 67 person team.
– Developed RAG component, data preprocessing, prompt engineering, text summarizing, model tuning.
– Design, develop and implement ML services, deploy to production on Bitbucket, GCP.
– Manage the product and collaborate with members of the support and engineering teams.
Tools: Flask, LLMs, GPT OpenAI API, NLTK, HuggingFace, BERT, Spacy, Converse.

ML ENGINEER — freelance   Oct 2021 - Nov 2021

– For forecasting voltage consumption developed a predictive Ml model(LSTM). Aquracy (RMSE) 81.7%.
– Developed prototype of web service for IoT anomaly detection with Autoencoder models.
Tools: Tensorflow, Streamlit, Anomaly detection, Time Series(TS) forecasting, LSTM, Autoencoders.

DATA SCIENTIST, ML ENGINEER — COMMUDATA   Feb 2021 - Oct 2021

NLP Startup (England, Portugal, Russia). The idea was - determine an author's psychological state from his texts.
– Data preprocessing, EDA, developed and checked hypotheses, visualization and dashboards creation.
– Developed text anomaly detection model and prototype of Log Anomaly Detection web-app.
– Improved the product library with new text-feature extraction functions, resulting in a 7% increase in ROC AUC.
Tools: Tensorflow, Gensim, Spacy, Autoencoders, LSTM, PCA/TSNE, TS, Streamlit, SQL.

LEAD CYBERSECURITY ENGINEER — PFR in the Komi Republic   Jun 2011 - Feb 2021

Regional dept. Pension Fund of Russia (160+ employees in regional HQ, 22 district branches)
– Experienced in collaborating with IT depts., stakeholders, as dept.head and supervised district depts.
– Deployed CyberSecSystems for compliance, developed Anti-FRAUD programs, created role model reducing access time in 10 times and extended access mgmt. system to other regional PRF offices.

Tools: VPN, Linux, bash, Cryptography, MS SQL.

PROJECTS

DariaAI   **As Sr.ML engineer**

Navigation chatbot for one of the top exhibition spaces globally (30M+ yearly visitors) website and mobile app (200K+ users). Implemented entity extraction, optimized search engine, and developed search logging and data prep service (annotation, summarizing) using LLMs, enhancing chatbot performance and UX.

AI Solutions for Analytics   **As Product manager, LLM engineer**

Developing LLM agents as database co-pilot, Ad-hoc analytics using Azure OpenAI API and LangChain.

Matching assortment for e-commerce   **As ML engineer**

Completed clustering car-parts based on technical fits and consumer characteristics. Developed the algorithm, prepared the dashboard. This enabled the following two projects.

Data coverage for product marketing team   **As Product manager, DS**

Developed a multi-dimensional algorithm for price segmentation and assessments. Led a team of analysts and ML engineers to design, implement, and use a data governance system for product marketing.
– Enabled analytics dept. to achieve complex goals unattainable for 18 months.
– Cut product launch prep time from 2 weeks to 1 day for product managers.

VIO-coverage   **As Data Scientist**

An assortment coverage gap assessment system. Matching big data of auto parts clustering and market demand. Updated heatmap of assortment gaps, boosting revenue by 8.7% yearly.

EDUCATION

2011 - 2016	PhD (Modeling and Security) at St. Petersburg Institute for Informatics and Automation Russian Academy of Sciences (SPIIRAS)
2005 - 2010	Masters Degree (with Honors) at Syktyvkar State University ranked top1 in the Komi Republic, Information security specialist

PUBLICATIONS

SKILLS

NLP	LLM, Transformers, GPT, LangChain, BM25, FAISS, Hugging Face, BERT, Spacy, Gensim, Glove, NLTK, prompt engineering, fine-tuning.
Data science	DL, Tensorflow, PyTorch, Prophet, Catboost, Autoencoders, LSTM, PCA, TSNE, Matching, Clustering, Anomaly detection, Time Series, Visualization.
Data	SQL (PostgreSQL, Clickhouse, MS SQL, MySQL), Metabase, BigQuery, Airflow.
Develpoment	FastAPI, Flask, Streamlit, Sentry, deploying on GCP, Bitbucket, GitLab ci-cd
Base	Python, Git, Docker, Linux, bash, Jira.