

# Alexander Borochkin

Data Scientist

## Contact

### Address

Oakland, CA 97232

### Phone

(971) 447 6676

### E-mail

aabor.nn.com@gmail.com

## Websites, Portfolios, Profiles

- <https://www.aabor.ninja>
- <https://www.linkedin.com/in/aaborochkin/>

## Skills

Python (pandas, numpy, scipy, sklearn, sklearn-onnx, matplotlib, keras, tensorflow, tensorrt, numba, pytorch)

SQL

R (dplyr, ggplot2, tidyr, dplyr, shiny, caret, tidyverse, e1071, plotly, knitr, xgboost)

AWS Solutions Architect

Neural Networks

Machine Learning

Natural Language Processing

Extensive experience in Python and R data science packages while working in large multinational companies as well as startups doing tasks 1-3 months long in small teams and taking lead eventually. Full stack developer of Artificial Intelligence products on different roles mostly on back end and analytics sides. Familiarity with cloud technologies, on-premises cloud, desktop virtualization. Knowledge in relational and non-relational databases, extract transform load or map reduce data processing. Design of high performance and resilient cloud architectures for big data processing in streaming, online or offline modes. Two international patents for Electronic Voice Assistant development as a team member.

## Work History

2022-10 -

2023-02

### Data Scientist

*Lad24, Nizhny Novgorod, Russia*

- In two weeks developed KPI Reporting component for Electronic Voice Assistant product.
- Helped company to migrate from MQ message broker to Apache Kafka when processing end user request Electronic Voice Assistant project. This allowed further workflow decoupling on NLP component level, increased product resilience and high availability. Decoupled NLP components decreased 3-5 times in size, conflicting dependencies on third party packages vanished.
- Designed automatic realtime KPI reporting system based on Apache Kafka and Apache Spark.
- Organized computation workflow in company cloud using Jupyter Notebooks on remote server and trained other Data Scientists to use them. This allowed many Data Scientists to use limited GPU resources in company cloud effectively.
- Worked as Solutions Architect consultant to DevOps team on AWS features.

2019-04 -

2022-10

### Data Scientist

*Harman Connected Services, Nizhny Novgorod, Russia*

- International patent reward for setup automated A/B testing framework for Multilingual Electronic Voice Assistant project and created automatic KPI

## Statistics

Linux, Docker, QEMU/KVM,  
Vagrant, Virtual Box,  
Parallels, Portainer

Terraform, Packer

Kafka, Spark,  
MongoDB/DynamoDB,  
Redis/ ElastiCache

Git, Gerrit, Jira, Bugzilla,  
Gitlab

XLST, HTML, Markdown,  
LaTeX

## Languages

English

Russian

Italian

Spanish

reporting component which were used by development team and stakeholders to coordinate efforts on new Natural Language Processing (NLP) and Speech to Text (STT) features development as well as to document expenses. Voice assistant performance increased from 60% to 90%.

- Created Extract Transform Load pipeline to decouple online transaction processing (OLTP) database from online analytical processing (OLAP) database preserving end user data privacy, feature unloaded compute heavy analytical read workflow from heavy but simple write workflow.
- International patent reward for semi automatic Dataset labeling methodology to train Neural Networks NLP and STT models for Electronic Voice Assistant Component.
- Different versions of PDF reports were made in R/Latex and Python languages.
- Developed algorithms for Natural Language Processing (NLP) team to expand, check and report statistics for multilingual training datasets. Algorithms can identify 2-3 mistakes in datasets that consist of millions of records saving hours of works of dataset engineer.
- Developed computer language to low code Voice Assistant behavior and logical choices as required from NLP Dataset team.
- Developed Punctuation and Capitalization Restoration Neural Network models in Russian and English languages for real time speech transcription and methodology for its performance estimation and comparison which showed performance results between major competitors like Skype, Zoom, Google.
- Participated in corporate events, took second place in corporate hackathon, developing mobile application for orientation inside buildings where GPS signal is not available.
- Utilized advanced querying, visualization and analytics tools to analyze and process complex NLP data sets.
- Improved data collection methods by designing

surveys, polls and other instruments.

2003-01 -  
2019-01

## Associate Professor

*State University of Nizhny Novgorod, Nizhny Novgorod, Russia*

- Published more than 20 papers in top Russian science Economics and Finance journals.
- Participated in international science conferences in Singapore, Buenos Aires (Argentina) among others, traveled frequently with students to conferences in Moscow, capital city of Russia.
- Created more than 5 curricula, lecture plans and slides for courses in Economics and Finance.
- Delivered more than 800 lecture hours per year, including traveling regionally.
- Organized yearly students contests between 5 regional universities.
- Edited student's research paper almanac for 6 years period.

---

## Education

---

2003-02 -  
2010-04

### Post graduate Degree: Banking, Corporate, Finance, Securities Law

*State University of Nizhny Novgorod - Nizhny Novgorod, Russia*

2001-02 -  
2002-12

### PhD: Business Administration and Management

*Università Della Calabria - Rende, Italy*

1997-09 -  
2002-01

### PhD: Banking, Corporate, Finance, Securities Law

*State University of Nizhny Novgorod - Nizhny Novgorod, Russia*

---

## Certifications

---

2023-09

[AWS Certified Solutions Architect - Associate], [AWS] - [3 months]

2019-04	[Big Data Analysis with Scala and Spark], [Coursera] - [3 months]
2019-04	[Functional Programming Principles in Scala], [Coursera] - [3 months]
2019-04	[Algorithmic Toolbox], [Coursera] - [3 months]
2018-02	[Data Scientist with R], [DataCamp] - [96 hours]
2018-02	[Data Scientist with Python], [DataCamp] - [84 hours]
2017-12	[Quantitative Analyst with R], [DataCamp] - [65 hours]