

Software Engineer, Research

Ilia Timofeev

ilia.timofeev@gmail.com

+49 1523 666 37 93

german work permit

github.com/iliatimofeev/

Python 3 year

C++ 7 years

Experience

Research lead, STC-innovations

August 2019 – present

STC is a national leader of speech analytics and biometry software, with 600 employees

More than 20 years of word recognized research in the field, latest achievements:

- First place at NIST SRE 2021 Track 1 (text-independent speaker recognition)
- First place at CHiME 2020 Track 2 (speech diarization and recognition)

Project: Large meeting transcription system (Python, Javascript)

- Boosted team performance at 72% (experiments per week)
 - Designed common experiments pipeline
 - Introduced system quality measurement as a part of continuous integration
- Grew team effectiveness at 60% (quality growth per month)
 - Introduced EDA and hypothesis prioritization framework at every stage of project
 - Implemented a standard set of data visualization tools
- Increased system quality by 38% in absolute (metric SA-WER)
 - Developed and executed an 18-month research roadmap with incremental quality growth thus finally having met customer expectations
 - Created dataset collection methodology to record 100 hours of acted meetings
 - Architected and supervised the implementation of microservice architecture consisting of:
 - Adopted for business domain In-house Automatic Speech Recognition SDK
 - In-house Face Recognition SDK
 - Diarization and speaker identification with Voice biometry models (Resnet)
 - Fine-tuned for conference room Silhouette Tracking model (FairMOT)
 - Zonal voice activity detector (xgboost)
 - Natural language processor for correction of speaker and ASR errors (BERT)
 - A sophisticated multistage decision algorithm
 - Infrastructure: Kafka, Docker

Project: Voice emotion recognition SDK for call center analytics software (Python, C++)

- Improved supervisor performance on monitoring task at 10% by conflict time identification
- Achieved 41% model quality (metric 4 class F1 – Score), presumable SotA at the moment
 - OpenL3 (VGGsh) embedding adoption and tuning for prosodic features
 - LSTM head trained on multitasking for hard and soft labeling approaches
 - Language model (XLNet) for sentiment analyses
- Patent requested

Data Science Lead, AdTech startup

December 2018 – June 2019

- Launched MVP within three months that allowed collection of transaction data (**Python**)
- 4 Data Insights reports delivered that allowed to reduce the cost of Ad campaigns by 2%

- Consultant, TTS Consulting (small consulting startup) 2015 – 2018
- Successfully executed international project in marketing and financial modeling
 - Defined transformation strategy that returned profitability for a client
- Head of Project planning, Roing (big national engineering company) 2011 – 2015
- Improved profitability by 20% by developing financial forecasting and controlling system
- Head of Project Management Office, Severstal 2010 – 2011
- Reduced projects delays by 30% by automating ongoing control and resource balancing
- Head of Process Engineering Group, Bercut (IT solution for telecom) 2007 – 2009
- Reduced annual budget on test hardware by 15% by implementing shared test pool
- Software Engineer, Skyros** 2001 – 2007

Skyros a national market leader of video security systems

Project: **Safe City CCTV** solution (C++)

- Designed and supervised the development of large scale system (installed in 5 cities)

Project: **Video storage** subsystem (C++)

- Developed multimedia storage that 2x faster than previous on large volumes
- Designed subsystem that allows replacing old data with new in a cycle without degradation
- Subsystem sold and installed on more than 1000 sites

Open Source Projects

- Contributor: [Altair](#) (Declarative Visualization in Python)
- Author: [GpdVega](#) (geospatial extension for Altair)

Conference Speaker

2019 DataFest6 ([video](#)) Altair: Interactive chats from Notebook to Web

2019 PyData#6 ([video](#)) Interactive chats in EDA with examples on Altair

Publication: [Altair: Interactive Statistical Visualizations for Python](#). JOSS

Education

2017 Machine learning and Data-analysis, MIPT and Yandex (Coursera)

2016 Data Science, Johns Hopkins University (Coursera)

2015 Mining Massive Datasets, Stanford University (Coursera)

2012 Team management and Leadership, Corporate education program

2010 Project Portfolio Management, PM University

2008 Project management on ANSI PMI PMBOK® 4rd Edition, PM Expert

2006 Project Estimation, Carnegie Mellon/TEKAMA

2006 Process and lifecycle of software engineering, RUSSEE (TEKAMA)

2006 Human factor in software engineering, RUSSEE (TEKAMA)

2005 Negotiation technology, Stockholm School of Economics

2005 Visual modeling, UML, OOA, IBM Rational Rose, IT Academy

2002 - 2005 **postgraduate course in Computer science**, Saint Petersburg State University

1997 - 2002 **MS in Mathematics**, Saint Petersburg State University

Tools

tensorflow, pandas, numpy, sklearn, scipy, xgboost, lightgbm, shap, pytorch, keras, matplotlib, seaborn, altair, postgresql, mongo, sqlite, grpc, rest, docker