Kononova Valentina

Data Scientist

Kazan, Russia
+7 927 4111 227
vvkononova1@gmail.com

Open-minded, team-oriented and result-focused data scientist, experienced in time series analysis and natural language processing using advanced ML algorithms.

EXPERIENCE

Reputation Company, Kazan — *Data Scientist*

June 2019 - present

- Development of a system which generates a relevant response, based on sentiment and semantic of text blocks
- Applying machine learning algorithms for text segmentation (spacy, nltk), clustering and classification (SVM, KNN, Naive Bayes, CNN, LSTM)
- Extraction of different features such as statistic data, NER, noun and verb chunks (spacy, nltk, nltk, google NL API)
- Planning and evaluation

Siemens Scholarship

September 2016 – August 2018

- Conducted researches on sensor data validation using ML algorithms (time series classification, clustering, visualization)

State Institute of Applied Optics, Kazan — *Engineer*

March 2016 – September 2017

- Designed documentation (the circuit diagram, Bill-Of-Materials)
- Provided quality assurance of printed circuit-boards and their components

EDUCATION

TU Ilmenau, Ilmenau — *master of science in computer science*

2016 – 2018

Kazan National Research Technical University named after A.N.Tupolev – KAI, Kazan — *master of science in computer science*

2016 – 2018

Kazan National Research Technical University named after A.N.Tupolev – KAI, Kazan — *bachelor in computer science*

2013 – 2016

SKILLS

Natural language processing
Time-series analysis
Machine Learning
Neural Networks

Python, SQL, Git, Docker, Kanban, Asana, Visual Studio, Atom, Jupyter Notebook

LANGUAGES

Russian	Native Speaker
English	Upper Intermediate
German	Elementary (A1)

INTERESTS

Personal Development

Foreign languages

Traveling

PROJECTS

Project's Name	Languages	Stack of Technologies
Text blocks clustering using different ML algorithms	Python	pandas, numpy, spacy, sklearn, SVM, kNN, Naive Bayes, LSTM, CNN, natural language processing, text embedding, BERT
Word sequence prediction	Python	pandas, numpy, keras, tensorflow, RNN, nlp
Text segmentation	Python	pandas, numpy, spacy, sklearn, natural language processing, syntax analysis, named entity recognition
User's feedback sentiment analysis	Python	pandas, numpy, matplotlib, nltk, keras, tensorflow, lstm, natural language processing
Determination of equipment operation modes	Python	pandas, numpy, matplotlib, sklearn, time-series analysis, statistics, lag window, Gaussian Mixture Model
High dimensional data visualization using PCA and t-SNE	Python	pandas, numpy, matplotlib, sklearn, edge detection
Object recognition using CNN	Python	numpy, keras,tensorflow, CNN, data mining
Sensor data validation	Python	pandas, numpy, matplotlib, sklearn, time-series analysis, features extraction, statistics, k-means, dbscan, hdbscan
Automating the placement of components of the printed circuit board in accordance with EMC using a genetic algorithm (Bachelor Project)	C#	SQLite, SQL, Unity, Visual Studio, OOP, genetic algorithm
An automated information system "Performances"	C#	MS Access, SQL, Visual Studio, OOP