

ALENA SOROKINA

CONTACT INFORMATION

E-mail: alena.sorokina@nu.edu.kz
Address: 53, Kabanbay batyr ave

Phone: +77057617440
Nur-Sultan, Kazakhstan

RESEARCH INTERESTS

Data Science, Machine Learning and its applications, Natural Language Processing

EDUCATION

Bachelor of Science, **Nazarbayev University** August 2015 - May 2019
Major: **Mathematics**, minor: Economics, GPA: 3.57

Foundation Certificate, **University College London** September 2014 - June 2015
Math and Physics, passed with Distinction

PUBLICATIONS

1. **A. Sorokina**, A. Karipbayeva and Zh. Assylbekov, "Low Rank Approximation of Matrices for PMI-based Word Embeddings". Springer "Lecture Notes in Computer Science". *International Conference on Computational Linguistics and Intelligent Text Processing (CICLing)*, 2019, **Oral, top 10%**
 - ArXiv pre-print: <https://arxiv.org/abs/1909.09855>
 - Implementation: <https://github.com/the-regressionists/low-rank-approx-for-PMI>
2. A. Karipbayeva, **A. Sorokina**, and Zh. Assylbekov, "A Critique of the Smooth Inverse Frequency Sentence Embeddings". *The Thirty-Fourth AAAI Conference on Artificial Intelligence, Student Abstract (AAAI 2020)*.
 - ArXiv pre-print: <https://arxiv.org/abs/1909.13494>
 - Implementation: <https://github.com/the-regressionists/Critique-SIF-Sent-Embeddings>

RESEARCH AND WORK EXPERIENCE

NLP/Speech researcher • Caption Call • Salt Lake City, Nur-Sultan September 2019 - current time

- Research on application of the neural networks to the audio super resolution problem

Research Assistant • Nazarbayev University • Department of Mathematics April 2018 - current time

- Supervisor: **Professor Zhenisbek Assylbekov**
- Research on the development of computationally efficient and interpretable algorithms for obtaining word embeddings and sentence embeddings.
- Working in Python (Numpy, Pandas, Scipy, Matplotlib, TensorFlow, Scikit-learn).

Research Intern • Ulsan National Institute of Science and Technology, South Korea Summer 2018

- Supervisor: Professor Bongsoo Jang's, School of Natural Science
- Research on applying networks theory for obtaining word embeddings

SKILLS AND LANGUAGES

Technical skills:

OS Platforms: Linux/Unix, Windows

Version control: ".git"

IDE: Visual Studio, Eclipse, Net-Beans, Oracle

Programming: Python, SAS, R, C, Java, SQL, L^AT_EX

Related Coursework:

- Statistical Programming, Machine Learning, Regression Analysis, Design of Experiments, Performance and Data Structures, Linear Algebra with Applications, Nonlinear Optimization, Algorithms

Languages:

- English – Fluent, Russian – Fluent, Kazakh – Upper-Intermediate, Korean – Beginner

OTHER EXPERIENCE

High school students mentor

Sep 2015-May 2019

- Helped students to set ambitious, yet realistic goals and work towards achieving them by familiarizing them with effective learning strategies and time-management skills
- Assisted with the university application process and entrance exams preparation (SAT, IELTS)
- Provided carrier orientation for high school students

International Hackaton "HackNU"

March 2019

- Solved Google task related to the CV and NLP

Bloomberg Market Concepts Certificate

November 2018

SELECTED AWARDS

- Full Undergraduate Scholarship
- Dean's List: Fall 2017, Spring 2017, Spring 2019
- Republic Chemistry Olympiad 2013 – Silver Medal
- Republic Chemistry Olympiad 2014 – Bronze Medal
- Altyn Belgi - School CGPA 5.0/5.0