

# ALENA SOROKINA

**Email:** alena.sorokina@nu.edu.kz, **Github:** alenasorokina, **Website:** alenasorokina.github.io

## RESEARCH INTERESTS

---

My research interests lie broadly in machine learning and its applications to natural language processing. I am particularly interested in probabilistic analysis of word embeddings.

## EDUCATION

---

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

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

## PUBLICATIONS

---

A. Karipbayeva\*, **A. Sorokina\*** and Zh. Assylbekov. “**A Critique of the Smooth Inverse Frequency Sentence Embeddings**”. In proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence, Student Abstract, **AAAI 2020**. *Github*

**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%**. *Github*

## RESEARCH AND WORK EXPERIENCE

---

**NLP/Speech researcher, Caption Call, Remote** *September 2019 - current time*

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

**Research Assistant, Nazarbayev University** *April 2018 - current time*

- Supervisor: **Professor Zhenisbek Assylbekov**, Department of Mathematics
- 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, UNIST, South Korea** *Summer 2018*

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

## TECHNICAL SKILLS AND LANGUAGES

---

**OS Platforms:** Linux, Windows      **IDE:** Visual Studio, Eclipse, Net-Beans, Oracle  
**Version control:** .git      **Programming:** Python, SAS, R, C, HTML, CSS, SQL, L<sup>A</sup>T<sub>E</sub>X

### 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
- Deans 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