Kamila Makhambetova

+1 (217) 550-2714 • kamila.makhambetova98@gmail.com • Github: https://github.com/kamilam3/kamilam3

EDUCATION

University of Illinois at Urbana-Champaign (UIUC)   Urbana-Champaign, IL
Master of Science in Statistics-Analytics, minor in Computer Science GPA: 4.0/4.0   May 2023

Selected Coursework: Fundamentals of Deep Learning, Natural Language Processing, Applied Bayesian Analysis, Statistical Learning, Statistical Modelling, Statistics and Probability II, Statistical Data Management, Data Structures and Algorithms in C++.
Award: Dr. John (Jack) McArdle *"Quantification Qualification"* Award for excellent data management and data analysis skills

Nazarbayev University   Astana, Kazakhstan
Bachelor of Science in Mathematics, GPA: 3.50/4.0   June 2021

Selected Coursework: Actuarial Science I, Regression Analysis, Linear Programming, Numerical Methods, Calculus, Linear Algebra, Differential Equations, Statistics and Probability, Stochastic Processes, Performance and Data Structures

SKILLS

Programming: Python (Numpy, Pandas, MatPlotLib, Scikit-learn, CPLEX, Gurobi, PyTorch), R (tidyverse, car, olsrr, leaps, lawstat, MASS, readr, jsonlite, rvest), SQL, Excel (VBA, Vlookup, pivoting, graphs, conditional formatting), C/C++, Java, SPSS, MATLAB, Mplus.

Algorithms / Skills: Supervised and Unsupervised Learning, Deep Learning, Natural Language Processing, Data Structures, Statistical Inference, Data pipelines, Design of Experiments, A/B Testing, Data Management, Exploratory Data Analysis, Data collection.

Software: Linux, Version Control: Github and Docker, Azure DataBricks, SQL Server, PowerBI, Tableau, Qualtrics, Bloomberg, MS Office.

Interests / Hobbies: Finance (interested in pursuing CFA charter), Fixed Income, Equity, world history, swimming, reading, yoga, and learning new languages (Native: Russian, Kazakh, Advanced: English, Elementary: Mandarin, French).

WORK EXPERIENCE

University of Illinois at Urbana-Champaign
Research Assistant at Mullen's Exercise, Technology, and Cognition (ETC) Laboratory   August 2022 – now

- Employing ***Machine Learning classification and predictive algorithms*** using ***Python*** and ***R*** to analyze the interrelationships between physical activity self-regulation and cognitive health using precision behavioral medicine framework.
- Designing and analyzing controlled experiments using ***statistical methods***. Managing, collecting, cleaning and analyzing ***structured*** and ***unstructured datasets*** to test Dr. Mullen's hypotheses for ML projects and publications.
- **Published managed datasets**: https://osf.io/u8ze3/ **Published research paper**: https://osf.io/preprints/psyarxiv/48jpt

American Fidelity
Data Science Intern   June 2022 - August 2022

- Implemented ***K-mean clustering*** using OSHA dataset in ***Python***, visualized results using ***Principal Component Analysis (PCA)*** to identify what type of insurance should be provided to the clients. This saved 20 hours of work in week.
- Built ***data pipelines, cleaned, extracted and managed*** various datasets in ***Azure DataBricks (Python, SQL)***.
- Created two interactive ***PowerBI Dashboards*** designed to enable leadership to track clients' financial data, and to track workforce management and analytics of American Fidelity employees.

ACADEMIC PROJECTS

Language models, text classification, neural machine translation, dependency parsing (CS 447: NLP)   Fall 2023

- Built classifiers based on ***Convolutional Neural Networks*** and ***Recurrent Neural Networks*** to detect the sentiment of movie reviews using the IMDb movie reviews dataset. Implemented ***encoder, decoder*** and trained sequence to sequence models for Spanish to English translation.

Face mask detection using AlexNet (STAT 430: Fundamentals of Deep Learning)   Spring 2022

- Based on Masked Face-Net dataset, implemented ***AlexNet from scratch in Python*** to build classification model that sorts photos into three groups: properly wearing a mask, improperly wearing a mask, and not wearing a mask. Created a program using **FaceNet face detection tool and trained AlexNet** that determines in real time whether a person is wearing a mask properly to monitor compliance with governmental health policy and sanitary rules to prevent the spread of COVID-19.

Predicting class labels using Fashion MNIST (STAT 542: Statistical Learning)   Spring 2022

- Implemented ***Self-Organized Maps, Random Forest, linear discriminant analysis, logistic regression, Gradient Boosting,*** and ***k-nearest neighbors*** algorithms in ***Python*** and assessed the models based on the maximum testing accuracy to classify Fashion MNIST data into the correct label classification.

Implementation of shortest path algorithms (CS 225: Data Structures and Algorithms Analysis)   Fall 2021

- The project was based on implementation of ***Dijkstra's algorithm*** and ***Floyd-Warshall Algorithm in C++*** to find the shortest distance path between two different cities. The graph class was constructed, where nodes were unique airport IDs, the edge was the presence of flights between two cities and the weight of the edge was the distance between these cities.