

Ksenia Lepikhina

Data Analyst and Statistical Modeler

Boulder, CO 80301
(303) 827-5501
kklepi@gmail.com
<https://klepikhina.github.io/>
[linkedin.com/in/ksenialepikhina](https://www.linkedin.com/in/ksenialepikhina)

EXPERIENCE

UCAR (GLOBE): Boulder, CO — *Software Engineer*

JAN 2020 - PRESENT

Led an AI/NLP project automating email responses (Python).
Used time series analysis to predict upcoming weeks ticket counts (Python, Slack).
Writing a white paper for the impact of COVID-19 on GLOBE data. (Python, Tableau)
Nominated for UCAR Outstanding Accomplishment Award 2020.
Represented the team in stakeholder meetings and provided data guidance across all teams.
Regularly queried SQL database to retrieve, modify, add, and delete GLOBE data.
Presented information about GLOBE data using visualization tools (Python, Tableau)
Led an intro to data science webinar for an international education audience.
Pushed for the modernization of communication by setting up Slack for our team.

Threat X: Louisville, CO — *Security Analyst*

JAN 2019 - NOV 2019

Researched how to predict the clustering of malicious entities (Python, Elasticsearch (ELK), Docker) and how to identify whether an entity is malicious using machine learning techniques (Python). Additionally wrote numerous scripts (Python, Bash).

The Trade Desk: Boulder, CO — *Data Science Intern*

MAY 2018 - JAN 2019

Optimized a fixed vs. variable deal classification problem within the TTD platform.
Built Machine Learning models (Decision Tree, Random Forest, and Logistic Regression) in Python to further identify and improve the classifications.
Rewrote pre-existing SQL queries to significantly optimize run time. Tested new queries to ensure faster run times.
Identified which models were most effective, and utilized S3 and Spark to hand off effective models for productionalization.

The Trade Desk: Boulder, CO — *Software Development Intern*

AUG 2017 - MAY 2018

Worked with VP of Innovations to build an automated survey for consumers in order to measure the success of retargeted ad campaigns (C#, S3)
Queried the database to identify correlation between ads served to consumers and their browsing history (SQL)
Actively worked in a large production code base, wrote logs, managed API's, and worked with Cloudberry S3 on a day-to-day basis.

CU Boulder (Earth Lab): Boulder, CO — *Data Science Intern*

JAN 2017 - AUG 2017

Explored the relationship between remotely sensed snow seasonality and active layer depth. (R) Studied the time series of annual permafrost temperatures. (Python)

EDUCATION

CU Boulder: Boulder, CO

M.S. Statistics and Data Science

AUG 2019 - MAY 2021

GPA: 4.0

B.S. Applied Math and Statistics

Minor: Computer Science

AUG 2015 - MAY 2019

COURSEWORK

Statistics:

Bayesian Statistics • Probability • Time Series • Spatial Statistics • Statistical Theory • Statistical Modeling

Applied Math:

Complex Analysis • Differential Equations • Data Assimilation • Fourier Series • Linear Algebra • Numerical Analysis • Real Analysis

Computer Science:

Algorithms • Computer Systems • Databases • Data Structures • Machine Learning • Network (Graph) Analysis

SKILLS

Programming:

Python • R • SQL • MongoDB • Spark • AWS • Tableau • MATLAB • LATEX • Mathematica • Bash • Google Analytics

OS:

Linux • Windows • Macintosh

General:

Agile Methodologies • Data Analysis • Machine Learning • Statistical Modeling

Other

Professional mountain biker (2015-2019)

Runner, skier, rock climber

Reader, painter