ELIZABETH KELESHIAN

keleshiane@gmail.com | /in/ekeleshian | /ekeleshian

TOOLS

LANGUAGES
Python
JavaScript
Postgresql

DATA
pandas
matplotlib
scikit-learn
numpy
nltk

FRONTEND
react
mobx

BACKEND
node
postgres

ML
tensorflow
fasttext
huggingface

OPEN-SOURCE WORK

COMMON VOICE (MOZILLA)
MARCH 2021 - PRESENT

Merged PR that adds over 25 MB of Armenian text extracted from wikipedia. Now leading crowdsourcing efforts for data collection through hosting meet-ups, making social media posts, applying for grants, and reaching out to academic institutions. Onboarded 17 volunteers, received over 10,000 impressions so far.

BIGSCIENCE (HUGGINGFACE)
MAY 2021 - PRESENT

Working with a team on personally identifiable information (PII) management, i.e. code to anonymize text scraped from the web in a scalable and secure way and be able to handle multiple languages. Using presidio for entity recognition and Helsinki-NLP for translations.

EXPERIENCE

DATA SCIENTIST
WAY2B1, INC. | SAN FRANCISCO | 05/19 - 01/21

Built Proprietorial Email Classification Chrome Extension
- Backend in node, flask and frontend in vanilla Javascript
- Offline training for text classification done with keras, preprocessing text with pandas, nltk, beautifulsoup, scikit-learn
- Model is a 1D ConvNet on word embeddings pretrained with Glove-300D
- Secured app with two layers of authorization using OAuth2.0

Built First Iteration of Development Support for Machine Learning
- Set up centralized logger and configuration for all microservices
- Built CLI's to handle decrypting text and extracting text from images/pdfs

Worked on Miscellaneous Frontend/Backend Features
- Came up with a new way to render long lists in the app using the react-virtualized library, speeding up loading by 7x
- Shipped over five major product features within four months

DATA SCIENCE FELLOW
DATA SCIENCE RETREAT | BERLIN | 09/18 - 12/18

Capstone Project: Recommendation System for Font Styles
Built a variational autoencoder to extract latent space of 5,000 unique font styles; utilized KNNs to determine recommended font types based on user's selection. Backend done with Google Cloud Functions, frontend done in D3.js

EDUCATION

UNIVERSITY OF CALIFORNIA, LOS ANGELES
B.S. MATHEMATICS/ECONOMICS, B.A. PHILOSOPHY | GPA 3.65

Relevant coursework includes Gaussian Statistics, Linear Algebra, Probability Theory, Discrete Mathematics, Optimization, Vector Calculus, Differential Equations, Stock Portfolio Risk Analysis (graduate level), Econometrics, Real Analysis

UNIVERSITY OF PENNSYLVANIA
M.S.Ed URBAN EDUCATION | GPA 3.5

Thesis: Investigating the effects from a pilot program providing free financial education to high school students living in underserved neighborhoods.