

COLLINS WESTNEDGE

✉ Cwestnedge@gmail.com ☎ 281-687-8052 📍 Chicago, IL in collins-westnedge/ 🚲 dirtylittledirtbike

SKILLS

PROGRAMMING: Python, Javascript, SQL, R, HTML

LIBRARIES: Pandas, Scikit-learn, Numpy, TensorFlow, nltk, Flask/Dash, spaCy, BeautifulSoup

DATA VISUALIZATION: Matplotlib, Seaborn, Plotly, Tableau

MACHINE LEARNING: Natural Language Processing, Regression, Classification, Clustering, Dimensionality Reduction

EXPERIENCE

Metis *Data Scientist*

Remote - Chicago, IL
Sept. 2020 - Dec. 2020

Attended a 12-week immersive data science bootcamp focused on project oriented learning. Completed self-designed data science projects from conception to presentation; including data collection, data management, exploratory data analysis, modeling, and visualizations.

r/depression (NLP, unsupervised)

Examined how mental health is discussed online as well as the benefits and common issues users face in these support spaces. BigQuery was used to gather posts over a 3 year period from the two largest and most active mental health subreddits, r/depression and r/anxiety. Dimensionality reduction techniques such as non-negative matrix factorization were used to ascertain common topics discussed in each subreddit. Time based analyses were used to assess the overall engagement and growth of these subreddits over time. Additionally, linear transformations were performed on learned word embeddings to shed light on the kind of feedback users desire as well as semantic biases for particular words present in the data.

Covid Risk Assessment (web app)

Created an interactive web application for visualizing Covid-related risk for various group sizes in a particular county. Data is gathered daily from the New York times for county level information on covid cases. Risk is defined as the estimated probability that at least one person in a group of a particular size is infected with covid based on county level statistics, BLS estimates for population size, as well as research findings by MIT concerning estimations for the factor by underreporting affects the data. The goals were, not only to illustrate how risk is relative to population size, but also to visualize the non-linear relationship that exists between risk and group size. The end product allows users to select and compare risk as well as time based figures for selected counties. Organizations that have been using the app include, Citizens Bank, Lisman Studios and Wrona Dubois Law Firm.

Toxic Comments (NLP, Classification)

Trained a multinomial bayesian classifier to classify different types of toxic speech. Data was gathered from wikipedia comments as part of a kaggle competition. Examined the coefficients for the trained model as well as topics gathered through NMF to unveil and resolve biases against LGBTQ terms inherent in the data. Overall identifying coefficients for bias terms in the training data led to the largest improvement on the models performance the test set.

Predicting Housing Prices (Regression, web scraping)

Scraped information on properties pertaining to the greater Chicago area from Zillow. Leveraged BeautifulSoup to collect and parse data based on specific search queries from the Zillow website. Created clean and usable data consisting of 22 unique property features for 2600 listing in the greater Chicago area. Used multiple regression models to predict housing prices. Random forests and tree based regressors were preferred. Overall they outperformed linear models and were able to capture non-linear relationships between price and and geospatial features to make more robust predictions.

Hanover Investment Advisors *Statistical Modeling and Programming Analyst*

Chicago, IL
Sept. 2015 - Oct. 2017

Worked with commercial real estate data and provided statistical analysis and insights in R for presentations to U.S. and European pension funds, pension fund advisors, insurance companies, sovereign wealth funds as well as U.S and European banks and Singapore family office.

EDUCATION

University of Chicago
Bachelor of Arts Philosophy 2017

School of the Art Institute of Chicago
Continuing Studies Art and Tech 2019