

# LOUISA REILLY

✉ lousacreilly@gmail.com ☎ 509-592-0319 📍 Seattle, WA 🌐 louisareilly 🌐 geeselouise

## SKILLS

**PYTHON:** Numpy, Pandas, Matplotlib, NLTK, Sklearn, Keras, Tensorflow

**DATA ENGINEERING:** SQL, Spark

**SUPERVISED LEARNING:** Linear Regression, Classification, Ensemble Learning, Logistic Regression, Neural Networks

**UNSUPERVISED LEARNING:** Clustering, Topic Modeling, Dimensionality Reduction, Natural Language Processing

**DATA VISUALIZATION:** Tableau, Seaborn, Plotly

**WEB-SCRAPING:** Selenium, BeautifulSoup

## DATA SCIENCE PROJECTS

### Classy DnD: Detecting Multiclass Characters

- The overarching goal of this project was to build an interpretive classification model that discerns whether or not a character has multi-classed based on character sheet data.
- Built a Tableau dashboard to visualize and understand the nuances of the dataset.
- All categorical features were one-hot encoded, and some numerical features were normalized to character level. The algorithms investigated were Logistic Regression, Random Forest, and Naïve Bayes, with Random Forest performing the best with a 0.55 F-1 score and 92.6% accuracy.

### Data Engineering and Visualization of Seattle Public Library Checkouts

- The Checkouts by Title dataset from the Seattle Public Library is a massive dataset of ~38.5 million entries. The goal of this project is to engineer an end-to-end data storage, management, and visualization pipeline.
- SoQL queries via the Socrata API were executed to obtain specific subsets of the data. Spark managed large chunks of data for sampling and processing.
- All data was saved into a SQL database, and then exploratory data analysis was performed using Tableau. Some insights gained were how current events, Seattle's culture, and month affect checkout trends.

### DeCoreIdea: Interior Design meets Deep Learning

- Interior design styles tend to be very fluid and subjective. Thusly, designers and their clientele often have difficulty communicating their ideas/ vision. To help solve this problem, multiple convolutional neural network multi-classification models were built via Keras to classify pictures of home interiors into specific design style categories.
- Model training was performed on Google Cloud Platform with the web-scraped images stored in Cloud Storage.
- To improve performance, transfer learning was leveraged with pre-trained weights from the imagenet dataset.
- Neural net architectures utilized were VGG16, MobileNet, and ResNet50; all models had issues with overfitting, but performed better than the Logistic Regression model which was merely guessing the categories.

### NLP Topic Modeling with Past Jeopardy Questions

- The overarching goal of this project was to discern more generalizable topics, so future Jeopardy contestants know what to prioritize when preparing to audition or be on the show; the game show is notorious for emphasizing breadth vs depth.
- After using non-negative matrix factorization (NMF) for topic modeling, this became evident, as the model needed 100 topics to create some discernible categories. Even then, the top category was mostly more general knowledge with a focus on history and geography.
- After some exploratory data analysis, it revealed that history and geography were in the top 20 of categories. Interestingly, some of the most popular categories were a hodge-podge of questions from 'Potpourri' and 'Before and After'.

### Targeting Truck Owners for Electric Vehicle Adoption

- The purpose of this project was to develop a fully-scoped proposal to address Ford Motor Company's need to refine their electric vehicle strategy.
- Through exploratory data analysis via Tableau and Excel, maps of the U.S. were created highlighting underserved rural communities (e.g. poverty, car-independent, unemployment rate, etc.), states with larger proportions of registered truck owners, and current/ future EV charging stations.
- The proposed data science solution was to use clustering analysis to divide regions of the U.S into groups based on community need and interest in EV trucks and commercial vans.

## EDUCATION

University of Washington  
M.S. Chemistry 2020

Gonzaga University  
B.S. Biochemistry 2019

## EXPERIENCE

Metis · Seattle, WA  
Data Scientist · June 2021 to Sept. 2021

- Completed a 14-week 280 hour immersive data science and machine learning bootcamp featuring a project-centered approach to achieve expertise in data handling, statistical modeling, and machine learning.
- Designed and delivered results for 7 statistical modeling and machine learning projects on the following topics: Exploratory Data Analysis, Regression, Classification, Business, Data Engineering, Unsupervised Learning/ NLP, and Deep Learning.
- See project section for more details.

University of Washington · Seattle, WA  
Teaching Associate · Jan. 2021 to Mar. 2021

- Facilitated the distribution of student unknown data packets for analysis.
- Lectured students virtually on organic chemistry lab concepts, techniques, and safety via PowerPoint and videos.
- Demonstrated how to analyze spectra and interpret ambiguous results.
- Graded student lab reports attended weekly lab meetings, and hosted office hours for students.

University of Washington · Seattle, WA  
Research and Teaching Assistant · Sept. 2019 to Dec. 2020

- Leveraged background in computational chemistry to develop Python programming skills through online tutorials in about 2 months after joining a natural language processing project.
- Worked collaboratively to web-scrape multiple chemistry textbook websites and academic journals (Wiley and Elsevier).
- Independently created a fully functional word2vec package and tutorial for future use in a named-entity recognition (NER) neural network that recognizes chemical entities in academic literature.
- Mentored 4 undergraduates, 2 high school students, and a community college professor in their development of Python programming skills in web-scraping and data visualization.

Gonzaga University · Spokane, WA  
Research Assistant · June 2017 to May 2019

- Completed 2 steps in a organic synthetic pathway for making a lactate analog to use a PET tracer in the brain to detect neurodegenerative diseases, such as Alzheimer's or Parkinson's.
- Finished the synthesis and analysis of medicinally significant isoprenoids attached to glucose to aid in the development of an assay for monitoring their metabolism in the human body.
- Synthesized and purified medicinally significant small organic molecules in anhydrous conditions via extraction and HPLC.
- Evaluated reaction results via 1 H NMR, 13 C NMR, IR spectroscopy, and GC-MS.

Gonzaga University · Spokane, WA  
Learning Assistant · May 2018 to May 2019

- Earned the International Tutor Training Program Level 1 Certification.
- Facilitated one-on-one tutoring sessions with 1-2 students in gen chem, ochem, ichem, and intro bio.
- Counseled students on good study habits and study skill development using SMART goals, the study cycle, and time management worksheets.
- Identified gaps in student knowledge by utilizing active listening skills, effective question asking strategies, and Bloom's Taxonomy.