

## EDUCATION

**University of Illinois at Urbana-Champaign** Urbana, IL  
Masters of Science in Statistics (GPA: 4/4) with specialization in Analytics Jan 2021 - Dec 2022

- **Coursework:** Applied Regression, Statistical Learning, Stats & Prob II, Algorithms for Data Analytics, Mathematical Statistics
- **Teaching Experience:** Graduate Teaching Assistant, course: CS 598- Foundations of Data Curation

**College of Engineering, University of Pune** Pune, India  
Bachelor of Technology in Mechanical Engineering (GPA: 8.6/10) Aug 2014 - May 2017

- **Relevant Coursework:** Linear Algebra, Numerical Methods, Optimization, Calculus, Statistics, Probability, Hypothesis Testing

## EXPERIENCE

**University of Illinois at Urbana-Champaign** Urbana, IL  
*Research Assistant at Grainger College of Engineering* Feb 2021 - July 2021

- Established a full stack web application to predict risk of railway track failure using **Python** and **Django** framework
- Automated data collection, analysis and prediction process using Python, cutting down manual efforts by 99%

**Eaton Research Labs** Magarpatta City, Pune  
*Data Scientist* May 2019 - Dec 2020

- Applied unsupervised learning to identify patterns in 20k engineering drawing files, implemented a **Support Vector Machines** model to classify drawing templates
- Set up an end-to-end pipeline using **Python**, fetching input files from PLM databases to extracting Material and Finish information utilizing **Google's Tesseract Engine** for OCR with 94% accuracy resulting in an annual savings of **\$0.5M**
- Created labeled dataset of 12k samples, cleaned text using **NLTK**, **Regex** to build supervised learning model for classification of incoming email type using SVM
- Established pipeline to generate email response as per class with 96% accuracy and min. F1 score of 85% for all classes
- Developed **Python** software with graphical user interface using **Tkinter** realizing in estimated annual savings of **\$0.1M**

*Software Development Engineer* Dec 2017 - Apr 2019

- Deployed full stack web applications using **Django** at backend, **JavaScript**, **CSS**, **HTML** at frontend
- Designed database models with relational **SQL** for environmental compliance request management as per requirements
- Automated manual processes using **Python**, **JavaScript** to increase efficiency, created an interactive user interface with standalone software, achieved **\$0.4M** annual savings

## PROJECT HIGHLIGHTS

*Top skills employers look for- Repository* July 2021 - Aug 2021

- Crawled over 15000 job postings from Amazon Jobs and Facebook Jobs for different job categories using **Python** and **Selenium**, cleaned and processed raw text data to identify top 20 skills for each job category
- Built an interactive dashboard using **plotly dash** with wordcloud, plotly charts, table; deployed on AWS

*Effects of diet on boosting immunity against Covid-19- Report* Feb 2021 - May 2021

- Performed analysis data from 165 countries, variable selection, modeled **parametric** and **non-parametric regression** algorithms with cross-validated RMSE of 0.035 using **Extreme Gradient Boosting Machines** algorithm in **R**

*Credit Card Fraud Detection- Report* Mar 2021 - May 2021

- Analyzed European credit card data from Kaggle containing 284k transactions to detect fraud; performed variable, model selection using **R**; Assessed **Logistic**, **KNN**, **Decision Tree** models, achieved C-V ROC curve area of 99.73 with **Xgboost** algorithm

## SKILLS

- **Programming:** Python (Pandas, Numpy, Tensorflow, Keras, Sklearn, Matplotlib, Plotly, etc.), R, SQL, JavaScript, VBA
- **Database, Visualization, Web Development Technologies** MySQL, Plotly, Power BI, Tableau, Django, VanillaJS, HTML, CSS
- **Other Tools:** Pycharm, Jupyter Notebook, VS Code, Jira, Miro, GIT, Slack, Microsoft Office

## OTHER ACHIEVEMENTS

- Received 3 employee of the month awards, 6 E-stars, recognition for mentoring a team of 5 members from project leaders, stakeholders at the innovation center, Pune Dec 2017 - Dec 2020