Jaskaran Singh Kohli

jkohli@scu.edu | 669-247-9973 | LinkedIn | GitHub | Palo Alto, CA

EDUCATION

Santa Clara University, Leavey School of Business — **Santa Clara, USA**

Master of Science in Business Analytics — **December 2020**

- Relevant Coursework: Machine Learning, Deep Learning, R and Python for Data Science, Natural Language Processing

SRM Institute of Science and Technology — **Chennai, India**

Bachelor in Computer Science and Engineering — **May 2018**

- Relevant Coursework: DBMS, Data Mining, Business Environment, Text Mining, Calculus, Statistics and Probability

TECHNICAL SKILLS

- **Data Analytics and Modeling:** Statistical Analysis, Hypothesis Testing, Data Mining, Data Wrangling, Feature Engineering, Marketing Analytics, Econometrics, Machine Learning, Supervised Learning (Regression, Classification, Ensembles), Unsupervised Learning (Clustering, PCA), Deep Learning (CNN,RNN), Natural Language Processing
- **Programming Languages and Tools:** Python, R Programming, SQL, Excel, SAS, Power BI, DAX, Power Query, Tableau, Oracle Analytics
- **Libraries and Frameworks:** Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, NLTK, Keras, TensorFlow, Dplyr, Caret, Ggplot2
- **Certifications:** Tableau Data Scientist Badge, Google Analytics Individual Certification, AWS Data Analytics Fundamentals, Project Management for Beginners, SAS® Academy for Data Science Badges, Master SQL for Data Science, Neural Networks and Deep Learning

EXPERIENCE

Gilead Sciences — **Foster City, CA**

Business Intelligence Analyst Intern — **June 2020 - September 2020**

- Streamlined robust data connections and automated manual reports for the Pharmaceutical Development and Manufacturing Department; reduced processing time by 60%.
- Applied different transformations for data conversion, sorting and cleaning using Power BI Query Editor. Developed calculations and visualizations using DAX functions such as table, aggregation and iteration functions. Produced a comprehensive business analysis and projections for future strategic decisions.
- Normalized the data and developed predictive analytics model by implementing Scikit-Learn pipelines to predict the number of days a list will take to reach the threshold and GridSearchCV to tune the hyperparameters; preventing crash during future website migrations.
- Optimized distribution, execution and scalability of live reports for 10 departments from 3 different data sources.
- Designed and deployed 30 ad hoc Power BI dashboards and built 15 Power BI applications for Dynamic Updates.

L&T Infotech — **Chennai, India**

Graduate Engineer Trainee — **August 2018 - January 2019**

- Analyzed and compiled several potential use cases for the business; delivered high performance technology solutions.
- Gathered client requirements, prepared multiple data analyses reports and conducted feasibility study for alignment with the Dev team.

Gaurish Technologies — **Gwalior, India**

Software Development Intern — **June 2016 - July 2016**

- Hands-on experience on a Live Project SCHOOL GURU'S School Management System deployed in more than 40 schools with data of more than 2500 students.
- Performed anomaly resolution and cataloging of errors concerning new services such as transcript and certificate generation with testing team; resulted in a better performing product for the end user and helped make error correcting process more efficient.

ACADEMIC PROJECTS

- **Oracle Practicum:** Collaborated with Oracle's Data Science & Advanced Analytics team. Conducted per product per company campaign analysis, feeding 450 Episodes with 50 states and 45 actions to implement Q Learning and build a marketing campaign recommendation system to help sales and marketing reps with the next steps aiming to increase the propensity of making a sale.
- **Neural Machine Translation using Transformers:** Implemented first sequence transduction model based entirely on attention, replacing slow recurrent layers with multi-headed self-attention, following the paper "Attention is all you need".
- **Stock Price Volatility Prediction:** Generated Glove embeddings from 17000 SEC 8K filings of S&P 500 companies from 2011-2020. Implemented several combinations of CNN, RNN and MLP Architectures concluded with accuracy of 64% to predict the stock volatility immediately after SEC 8K filings with CNN, RNN combination .
- **Recipe Generator using RNN** : Built a model by training 100k recipes using LSTM and RNN which will suggest an entire recipe along with cooking instructions and ingredients when the user inputs an ingredient.
- **US Unemployment Rate Time Series Analysis:** Analyzed the monthly U.S. unemployment rate containing 900K observations spanning over 27 years. Implemented STL decomposition and selected ARIMA model to fit the time series and forecast rate for the next 4 years.
- **Causal Effect of Cigarettes on CHD:** Analyzed several risk factors of developing a coronary heart disease to understand the causal relationship of cigarette smoked per day as a factor in determining whether a patient suffers from Coronary Heart Disease.