Kankshat Patel

kankshatp@gmail.com <> kankshat.github.io <> www.linkedin.com/in/kankshatpatel

EDUCATION

University of California, Santa Barbara • Bachelor of Science 2022 in *Data Science* and *Statistics*
Organizations: Data Science Club *Vice President*

SKILLS

- *Languages*: Python, R, SQL, C++, Java
- *Machine Learning Models*: KNN, Logistic/Linear Regression, PCA, Decision Tree, Random Forests
- *Elasticsearch*: Elasticsearch, Kibana, Logstash, ECK on Kubernetes.
- *DevOps:* Ansible, Git, Jenkins

PROFESSIONAL EXPERIENCE

University of California Santa Barbara Analytics Department **July 2021 – June 2022**
Data Analytics Intern

- Integrated KNN, PCA, Decision Tree, and Random Forest models in R and Python to precisely predict the outcome of a Division-1 batter's plate appearance with 80% success rate given various parameters.
- Engineered a data scraper in python using the Beautiful Soup module and regular expressions to extract and clean the data from the NCAA site transforming it into csv format, optimizing the time retrieval by 75%.
- Created ECK operator and installed Elasticsearch and Kibana environment through Kubernetes.

JP Technology **June 2021 – Sept 2021**
Software Developer Intern

- Installed and Configured Elasticsearch cluster through ansible script to streamline the process and reindexed data into new Elasticsearch cluster with zero downtime.
- Designed and implemented component templates to compartmentalize index templates into three smaller templates, simplifying the process of making changes to index settings and mapping for future index adjustments.
- Updated and installed an existing index extension within production which contained a txt file used by multiple index analyzers on Elastic Cloud Environment (ECE), increasing search capabilities and hits by 20%.

OhloneHacks **Nov 2019 - April 2020**
Founder/Company Recruitment Liaison

- Founded and organized colleges first ever hackathon, hosting 300+ participants.
- Actively recruited companies to sponsor hackathon, working closely on technology stacks available to participants.

PROJECTS

Relationship Compatibility Data Science Model **Aug 2020**

- Integrated ML algorithms with normalized and interpolated data to predict matches for speed dating.
- Built a classification model in Python using a k-means clustering method to predict finding a match.

COVID-19 Case Prediction Model **Dec 2021**

- Extracted relevant COVID-19 data from CovidCast API and tidied dataset to apply ML algorithms on time series analysis.
- Integrated Linear Regression and PCA and found a 77% success rate in predicting the number of new COVID-19 cases a week in advance.

RELEVANT COURSEWORK

Machine Learning and Algorithms || Statistical Machine Learning (Graduate) || Linear/Logistic Regression || Probability and Statistics || SAS Base Programming || Stochastic Processing || OOP (C++, Java) || Discrete Structures || Linear Algebra || Advanced Data Structures and Algorithms || Assembly Language