

# Ketan Patil

☎ (+1) 480-307-2774

✉ kkpatil1@asu.edu

🏠 ketanpatil.herokuapp.com

📱 Thek10patil

🌐 ketan-patil10

## Education

### Arizona State University

MASTERS IN COMPUTER SCIENCE, GPA = 3.85

Research Assistant in CIPS Lab

Intended Graduate- May 2020

- **Courses** - Distributed Database System, Multimedia web databases, Software Verification/Validation/Testing, Data Visualization, Statistical Machine Learning, Data Mining, Introduction to Artificial Intelligence, Foundation of Algorithms, Mobile Computing.

### College of Engineering, Pune

B.TECH IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY, RANKED IN TOP 10, GPA = 8.1 OUT OF 10

August 2012 - May 2016

- **Relevant Courses** :- Algorithms and Complexity, Operating System, Principles of Programming Language, Information Retrieval, Introduction to Business Analytic, Data Mining, Big Data and Cloud, Web Development, Calculus, Probability and Statistics.

## Skills

### PROGRAMMING LANGUAGE

- Python, Java, JavaScript, Android Application development, R, C, AngularJS, HTML.

### DATABASES

- Elasticsearch, AWS DynamoDB, MySQL, PL/SQL, SQLite, PostgreSQL.

### FRAMEWORKS/LIBRARIES

- AWS Glue, AWS Lambda, Hadoop, Spark, S3, Tableau, D3, Django, Flask

## Work Experience

### Amazon

SDE INTERN, DEVICE ORGANIZATION

May 2019 - August 2019

- **System architecture design & development** of Service Health dashboard is to view aggregated exception signature, aggregated error-level logging class names, high latency traffic of onboarded Alexa services on Atocha which handles **66 billion number of real time logs per day** using **AWS Elastic Load balancer, Elastic Search, S3, and DynamoDB technologies**. It reduced query time from 50-60 seconds to less than 10 seconds and increased search time window from 3 days to 30 days.
- Completed **Architecting on AWS featuring AWS labs** training offered by Amazon AWS.

### Credit Suisse

TECHNOLOGY ANALYST IN RISK AND FINANCE IT

July 2017 - July 2018

- Developed a summary email dashboard using text mining machine learning models built with **LDA, Bag of words model, and Contextual Analysis in python**. Created dashboard used for showcasing analytical view and efficiency calculation of around 10k employees helped the organization to bring down the manual monitoring efforts by **40%**.
- Designed a dashboard using **AngularJS** for business users with analysis of **70 TB** live market data in the graphical and tabular form which is useful for trading in the Market Risk Devision using Java language for backend development and MySQL database.

TECHNOLOGY ANALYST IN GROUP FINANCE DOMAIN

July 2016 - June 2017

- Analyzed batch performance and failure patterns with an aim to identify trends and causes in 10 years of historical PeopleSoft Batch data using Impala database with the help of **Random Forest for classification and Naive Bayes clustering algorithms**.
- **Communicated** new ideas and strategies over email to team of 32 members spread across different geographical locations.

### Credit Suisse

INTERN

May 2015 - July 2015

- Applied **text mining on Credit Risk market production data** for predicting Credit Score of relevant counter-parties using domain knowledge and analytical skills. Intern Member of **organizing team** who organized Classroom training and speaker series with higher management.

## Academic Projects

IMAGE SEARCH ENGINE. MULTIMEDIA WEB DATABASES

Fall 2019

- Developed a search engine dashboard to search similar images using various dimensionality reduction and feature extraction techniques from 11k image multimedia web database.
- Developed image classifiers using classification techniques like Support Vector Machine and Decision Tree, Personalised Page Rank
- Included relevance feedback functionality for improving query result based on user's inputs.

CASE STUDY: MANUFACTURING QUALITY COMMONALITY, INTEL

Winter 2018

- The assembly and test production model of manufacturing at Intel Corporation employs a form of batch processing. In batch processing, groups of units are processed together through different stations along the manufacturing process flow. The station that performs the final testing of the units has noticed some irregularities regarding the failure rate of the units.
- Analyzed parametric **highly imbalanced class data using sampling methods** to detect such a problem while processing the units thereby preventing such an increase in defects using different machine learning models like **XGBoost, Random Forest and Decision tree Classifiers**.

INSINCERE QUESTIONS CLASSIFICATION - COMPARATIVE STUDY OF ML MODELS, SML

Spring 2019

- Examined different ways and methodologies for better text pre-processing and detailed result comparison of Machine Learning algorithms implemented using graphs.

VISITOR MANAGEMENT ANALYSIS SYSTEM

Aug. 2015 - May 2016

- Worked on a **Django Python Web framework** to develop a dashboard with the incorporation of **support vector machine algorithm** to provide a better insight of hospitals, colleges and hotels visitors to organizations and present summarized data using statistical charts.