Chaitanya **Jadhav**

Machine Learning Engineer · Computer Science Engineer

Mumbai, Maharashtra, India

(+91) 9049044409 | chaitanya.jdh@gmail.com | www.chaitanyajadhav.com | Chaitanya-Jadhav | chaitanya-jdh

Career Objective

I possess an in-depth knowledge of Machine Learning, Natural Language Processing, Deep Learning, and various tools and techniques that are used in the Data Science/ Artificial Intelligence domain.I am a computer geek at heart and skilled in the programming field. I have succesfuly completed my end to end Master's thesis project, **Automatic Speech Recognition** at **State Bank of India**.

I would like to work in an organization that will give me a platform to utilize my problem solving skills and enrich my knowledge in the process to help growth in Machine Learning and Artificial Intelligence domain.

Education

School of Data Science and Forecasting, D.A.V.V. — *Indore, M.P.*

M. Tech in Data Science — *2018 - Present*

- **CGPA:** 8.94

N.B.N Sinhgad School of Engineering — *Pune,Maharashtra*

B.E. in Computer Science and Engineering — *2012 - 2016*

- Graduated with First Class

Skills

Programming	Python, JAVA, Shell Scripting, LaTeX
Machine Learning	Linear Regression, Logistic Regression, Decision trees, Random forest, Clustering (K-Means, Hierarchal), Gradient decent, SVM, Deep Learning (CNN, RNN, ANN) using TensorFlow (Keras).
Packages/libraries	Pandas, NumPy, Seaborn, SciPy, Matplotlib, Scikit-learn, BeautifulSoup
Tools	MS-Excel, Kaldi Toolkit
Database	MySQL, SQLite
Operating Systems	Linux, Windows
Server	Kaldi G-streamer
Familiar with	Git Version Control, Docker

Experience

State Bank of India — *Mumbai, India*

Speech Recognition Intern — *Aug. 2019 - Present*

- Provisioned an easily managable hybrid infrastructure(Amazon AWS + On-premise) utilizing IaC(Infrastructure as Code) tools like Ansible, Packer and Terraform.
- Built fully automated CI/CD pipelines on CircleCI for containerized applications using Docker, AWS ECR and Rancher.
- Designed an overall service architecture and pipelines of the Machine Learning based Fashion Tagging API SaaS product with the micro-services architecture.
- Implemented several API microservices in Node.js Koa and in the serverless AWS Lambda functions.
- Deployed a centralized logging environment(ELK, Filebeat, CloudWatch, S3) which gather log data from docker containers and AWS resources.
- Deployed a centralized monitoring environment(Grafana, InfluxDB, CollectD) which gather system metrics as well as docker run-time metrics.

Eran Group — *Indore, India*

Inventory Forecasting Intern — *May. 2019 - Jul. 2019*

- Implemented RESTful API server for car rental booking application(CARPLAT in Google Play).
- Built and deployed overall service infrastructure utilizing Docker container, CircleCI, and several AWS stack(Including EC2, ECS, Route 53, S3, CloudFront, RDS, ElastiCache, IAM), focusing on high-availability, fault tolerance, and auto-scaling.
- Developed an easy-to-use Payment module which connects to major PG(Payment Gateway) companies in Korea.

Remote Data e-Xchange — *Pune, India*

Software Developer — *Mar. 2018 - Jul. 2018*

- Researched classification algorithms(SVM, CNN) to improve accuracy of human exercise recognition with wearable device.
- Developed two TIZEN applications to collect sample data set and to recognize user exercise on SAMSUNG Gear S.

Projects

Automatic Speech Recognition System for Indian Accented English

Mumbai, India

State Bank of India

Aug. 2019 - Present

- Developed a Speech Recognition System for State Bank of India primararily targeted for Indian Accented English using **Kaldi Toolkit**.
- Achieved **14 percent WER** on banking test set.
- ASR system was trained on data collected from NPTEL lectures, subset of Commonvoice data comprising of Indian Speakers and Dataset obtained from I.I.T Madras.
- Used Beep Dictionary for lexicon and trained **G2P sequiter** model to produce pronunciations for OOV(Out of Vocabulary) words.
- Custom Language model was trained using **SRILM toolkit** over text corpus collected from various banking websites using Common Crawl.
- Hosted Speech Recognition Web Socket server for kaldi model using **Kaldi G-Streamer** and Android client.

Inventory Demand Forecasting

Indore, India

Eran Group

May. 2019 - Jul. 2019

- Developed a Time Series Forecasting model for Inventory Management.
- The primary task was to predict the order of SKU at RDC level inorder to maintain inventory to prevent over stocking and under stocking of various SKUs

PIDS Alarm Verification System

Pune, India

Remote Data e-Xchange

Mar. 2018 - Jul. 2018

- Developed a Verification System for Pipeline Intrusion Detection System Alarms.
- The fiber senses the Acoustic Vibrations that take place in its vicinity and generate alerts.
- Alerts consist of location details of the intrusion.
- These alerts must be visited by the Monitoring officers
- The verification system sents the newly generated alert details to the monitoring officer and marks the alert visited if the officer visits the alert location.

Emergency Android Application

Pune, India

Remote Data e-Xchange

Mar. 2018 - Jul. 2018

- Developed a module in which if a employee pressed a power button three times immediately a text mesage is sent to a supervisor/ concerned person.
- Text message consist of employee details along with GPS location.
- This application is used by the emplyees monitoring the petroleum pipeline.

Payroll Web Application

Pune, India

Remote Data e-Xchange

Mar. 2018 - Jul. 2018

- Built and developed a salary calculating web application which is intended to generate salaries of employee based on their designation, their class and location
- Application is enabled with feature to provide salary statement of all the employees in excel and per employee salary slip's in pdf
- Proposed various marketing and network activities to raise awareness.

Women's Safety Application using virtual Geo- fence

Pune, India

N.B.N Sinhgad School of Engineering (Academic Project)

Mar. 2018 - Jul. 2018

- Developed a Emergency Security Android Application which is primararily targetted for the use and safety of working womens.
- The application provides makes use of the GPS to track the mobile location.
- The application generate and send alerts to the users friends,family and to their respective company if the user crosses the virtual geo-fence around the route and does not return within geofence within specified time duration.

Self Assessment Assignments

T-SNE Visualisation on Amazon reviews with polarity based color coding

Applied AI course

May. 2019 - Jul. 2019

- Visualized User Review Text using T-SNE algorithm.
- Analyzed variation in T-SNE visualization on BOW, TFIDF, AVG-W2V and TFIDF-AVG W2V vectorizers.

K-NN classifier on Amazon reviews data set

Applied AI course

May. 2019 - Jul. 2019

- Built Brute force KNN and KD-tree version of KNN classifiers on BOW, TFIDF, AVG-W2V and TFIDF-AVG W2V vectorizers.
- Tuned hyper parameter K with N fold validation using AUC as performance metric
- Represented result for best value of hyperparameter K with ROC plot for train and test set along with confusion matrix for test set.
- Analyzed results for various combinations of vectorizers and versions of KNN for best hyper parameter.

Naive Bayes classifier on Amazon reviews

Indore, India

Applied AI course

May. 2019 - Jul. 2019

- Applied Multi-Nomial Naive Bayes classifiers on BOW, TFIDF, AVG-W2V and TFIDF-AVG W2V vectorizers.
- Tuned hyper parameter lambda using AUC as performance metric.
- Represented result for best value of hyperparameter alpha with ROC plot for train and test set along with confusion matrix for test set.
- Performed Pertubation testing to understand multicollinearity with one vectorizer.
- Also calculated Sparsity on Weight Vector obtained using L1 regularization for classifier built over one vectorizer.
- Listed top 10 important features for both positive and negative classes separately using weight vector of LR classifier.

Logistic Regression classifier on Amazon reviews data set

Pune, India

Applied AI course

Mar. 2018 - Jul. 2018

- Applied Logistic Regression classifiers on BOW, TFIDF vectorizers.
- Tuned hyper parameter alpha i.e Laplace smoothing factor using AUC as performance metric.
- Represented result for best value of hyperparameter alpha with ROC plot for train and test set along with confusion matrix for test set.
- Listed top 10 features (words) of positive reviews and top 10 features (words) of negative reviews using values of feature-log-prob- parameter of MultinomialNB.

SGD Implementation for linear regression

Pune, India

Applied AI course

Mar. 2018 - Jul. 2018

- Implemented Stochastic Gradient Descent for Linear Regression on Boston House Pricing Dataset.
- Compared the result with Scikit-Learn implementation of Linear Regression.

SVM on Amazon reviews data set

Pune, India

Applied AI course

Mar. 2018 - Jul. 2018

- Applied SVM classifier on BOW, TFIDF, AVG-W2V and TFIDF-AVG-W2V vectorizers.
- Worked with Linear Kernel and RBF Kernal version of SVM on all vectorizers
- Tuned hyper parameter alpha for both L1 and L2 regularization using AUC as performance metric.
- Represented result for best value of hyperparameter alpha with ROC plot for train and test set along with confusion matrix for test set.
- Listed top 10 important features for both positive and negative classes separately using weight vector for linear SVM.

Decision Trees on Amazon reviews data set

Pune, India

Applied AI course

Mar. 2018 - Jul. 2018

- Applied Decision Tree classifier on BOW, TFIDF, AVG-W2V and TFIDF-AVG W2V vectorizers.
- Tuned 2 hyper parameter best-depth and min-samples-split using AUC as performance metric.
- Represented result for best value of hyperparameter alpha with ROC plot for train and test set along with confusion matrix for test set.
- Visualized decision tree with Graphviz library to understand how a decision is being constructed given a new vector.
- Listed top 20 important features for both positive and negative classes separately feature-importances- method of Decision Tree Classifier.

K means Agglomerative DBSCAN clustering algorithms on Amazon reviews data set

Pune, India

Applied AI course

Mar. 2018 - Jul. 2018

- Applied K-means Clustering, Agglomerative Clustering and DBSCAN on BOW, TFIDF, AVG-W2V and TFIDF-AVG W2V vectorizers.
- For K Means Clustering found the best 'k' i.e the number of clusters using the elbow-knee method on k vs inertia- plot where inertia- is the Sum of squared distances of samples to their closest cluster center.
- Plotted Word Cloud for every clustering technique over various vectorizers to understand the nature of each cluster.

Examination

2018 **AIR 9000**, GATE Computer Science and Information Technology

Certification and Courses

Applied AI Course

Mumbai, India

Machine Learning Course

Aug. 2019 - Present

MindScripts Technology

Pune India

Diploma in Java A Grade

Aug. 2019 - Present