# Smit Patel

+91 97123 99167 | smitrpatel19@gmail.com | linkedin.com/in/jayambe | github.com/jayambe36

## Education

**Bachelor of Technology, Computer Science and Engineering** — **July 2019 - May 2023**
*Ganpat University, INDIA* — ***7.10/10***

**Big Data and Analytics** — **July 2019 - April 2023**
*IBM, Certificate Link*

**Machine Learning** — **Feb. 2021 - Sep 2021**
*Stanford University, Coursera Certificate Link* — ***94.5/100***

**AWS Academy Graduate-Machine Learning Foundation** — **Feb. 2022 - Sep 2022**
*AWS Academy Badge Link*

## Experience

**Software Engineer Trainee** — Jan 2023 – April 2023
*IBM - Career Education Program* — *Certificate Link*

- Created a Python-based IPR filing system using PostgreSQL, Docker, AWS, and Jenkins, enhancing efficiency and accuracy.
- Developed a Flask application with SIFT and SURF algorithms to facilitate IPR filing for the grassroots community.
- Explored ways to visualize GitHub collaboration in a classroom setting

**Machine learning Engineer** — May 2021 – Jun 2021
*Panache Software*

- Explored methods to detect object in live webcam and CCTV
- Developed a high-precision Face Mask Detection system using Python, CNN, TensorFlow, and Keras
- Contributed 500+ lines of code to an established system via Git
- Implemented real-time face capture and detection via CCTV and webcam, processing over 1000 frames per second
- Wrote 8-page ppt and gave multiple work-presentations
- Presented virtually to the World Conference on Computational Intelligence

## Technical Skills

**Languages**: Python, HTML5, Tailwind CSS, JavaScript, C++, C, R, SCALA
**Frameworks**: Django, Flask, ReactJS, NodeJS, ExpressJS, ReactNative
**Developer Tools**: GoogleColab Notebook, VSCODE, R-Studio Jupyter Notebook, Anaconda, Git , Razorpay, Cloudinary, Expo
**Tech**: OOPS, API, PaymentGatewaySystem, ImageVideoUploadSystem, SecureAuth (jsonwebtoken)
**Database**: MongoDB(Hybrid/Cloud connecting MERN Project, ML project with Databases), Relational Database - MySQL (Queries, Joins, Trigger, Function, Stored Procedures), DynamoDB, SQLite(connecting Python with Databases), IBM DB2
**HostPlatform**: AWS (EC2,ECR,S3,Sagemaker), Render, Vercel, Heroku, pythonanywhere, 000webhostapp, Godaddy, Shinnyapp, GithubPage
**IBM Tools**: IBM Watson, IBM SPSS Modeler, IBM CLOUD, IBM Cognos BI, IBM InfoSphere BigInsights, IBM CHATBOT, IBM DB2
**Data Science**: Predictive Modeling, Statistical Analysis, MLOps, Deep Learning, Data cleaning , Data Visualization(PowerBI and Tableu), Data Mining and Warehousing (Data Cube and OLAP, Schemas for multidimensional data models), Data Transformation, Data Standardization, Data Normalization, Handling Missing Data, Data Lake, JSON data format, ETL, Data Pipeline, Data Wrangling ,Data Summarization, Correlation Analysis, Grouping and Aggregating Data, BeautifulSoup and Scrapy(Web Scraping)
**ML/AI**: Supervised and Unsupervised, Dimensionality Reduction, Neural Networks (ANN, RNN, CNN), Natural Language Processing (NLP), Inferential Statistics, Association Rules, Classification and Prediction (K-means, Decision Tree, Bayesian classification, etc.), Clustering (KNN), EDA (Explorartory Data Analysis), ARIMA Models, Model Evaluation and Validation, Model Hyperparameter Tuning
**Big Data**: Pyspark, Scala, Hive, Hbase, Hadoop Ecosystem (HDFS, MapReduce, Yarn, PIG, HIVE, HBASE, Zookeeper), ApacheSpark, Statistical Analysis (Hypothesis Tests - Z-test, T-test, Chi-square Test), Probability and Statistics (Measure of central values, Measure of dispersion, Random Variables), Linear Algebra, Calculus

**NLP techniques**: BoW(Bag of Word), Word2Vec, Transformers, BERT, Sentiment Analysis
**Libraries/Algorithm**: A* Algo, Scikit-learn(Regression and Classification), Ensemble methods(Random Forest(Bagging) and Gradient Boosting, XGBoost (Boosting)), Tensorflow Keras (Deep Learning), LSTM( time series prediction, NLP, and speech recognition), NLTK(text processing for classification, tokenization, stemming, tagging, parsing, and semantic reasoning), Plotly and Matplotlib(creating static, interactive, and animated visualizations in Python, including line plots, scatter plots, bar charts, histograms), Seaborn(with Matplotlib: high-level interface for informative statistical graphics, whilst with Panda:visualizing univariate and bivariate data, as well as linear regression models and statistical time series), SciPy(scientific and technical computing), Pandas(data analysis and manipulation), NumPy(numerical computing)

## Coding Profiles

- LeetCode, GeeksForGeeks, AlgoZenith, HackerRank, CodeForces, CodingNinjas, Kaggle, CodeChef

## Badges

- Big Data: IBM, Machine Learning: AWS, Cyber Security: Cisco, CCNA: Cisco.
- Generative AI, LLM : Google Cloud Skill

## Personal, Academic and Freelance Projects

**Company Bankruptcy Prediction** | *Python, IBM SPSS Modeler*
- Led the development of a high-accuracy (99.045%) predictive model for company bankruptcy using a comprehensive financial dataset (97 columns, 6820 rows, 10 years)
- Implemented advanced data cleaning and feature selection techniques, ensuring robust data quality and meaningful financial insights for stakeholders
- Utilized machine learning models to analyze key financial indicators, enhancing decision-making processes for creditors and investors
- Conducted extensive data analysis and model validation, significantly improving the reliability and performance of financial risk assessments

**Employee Churn** | *Python,IBM SPSS Modeler*
- Led a 3-member team in creating a predictive model using SPSS Modeler and Python to evaluate candidate retention (0 or 1) utilizing the provided Training Data, a pivotal component of our project's success
- Addressed data quality and model accuracy challenges through thorough data cleaning, normalization, and reclassification, resulting in a 98% accuracy rate on the Test Data set, bolstering the model's reliability and effectiveness

**IBM Project-IPR filing for the Grassroots Community** | *Python, Docker, Ec2, Git*
- Engineered an IPR(Intellectual property rights) Filing System in Python with PostgreSQL, enhancing filing accuracy and operational efficiency
- Leveraged Docker for containerization, achieving consistent deployment across development and production environments.
- Integrated AWS EC2 for hosting and S3 for scalable data storage, optimizing cloud resource utilization
- Implemented Jenkins for CI/CD pipelines, automating testing and deployment processes for faster release cycles

**Panache Software Project-Face Mask Detection system** | *Python, Machine Learning Algo*
- Developed an advanced Face Mask Detection system using Python, leveraging CNN, TensorFlow, and Keras, achieving 95
- Integrated CCTV and webcam functionalities for live face capture and mask detection, processing over 1000 frames per second
- Conducted data analytics on 10,000+ captured faces to enhance the detection algorithm's precision
- Collaborated with a team of 3 to ensure robust performance and deployment of the detection system in industrial environments

**TEXT DATA MANIPULATION AND ANALYTICS** | *Python, NLP*
- Executed a Text Data Manipulation and Analytics project using Python, implementing the Bag of Words model for text visualization and NLP
- Conducted comprehensive text cleaning, including converting text to lowercase and word tokenization, processing over 50,000 text entries.

* Utilized word tagging (POS tagging) and analyzed tag information, enhancing the accuracy of text data categorization by 20
* Automated the counting of each word and tag information extraction, significantly improving the efficiency of text data processing workflows

**Champagne Sales Prediction** | *Python, Google Colab Notebook, Machine Learning Model/Algo*
* Analyzed 9 years of champagne sales data to forecast future sales trends using advanced statistical techniques
* Performed stationarity checks with rolling statistics and the Augmented Dickey-Fuller test, applying differencing to achieve stationarity
* Developed and compared ARIMA and SARIMA models, selecting the SARIMA model for optimal forecasting accuracy
* Successfully implemented the SARIMA model to predict future sales, enhancing decision-making processes with data-driven insights

**StudyNotion** | *MongoDB, ExpressJS, ReactJS, NodeJS, API*
* Engineered a fully functional ed-tech platform using the MERN stack (MongoDB, ExpressJS, ReactJS, NodeJS) to enable users to create, consume, and rate educational content
* Designed and implemented a responsive and interactive front end with ReactJS, enhancing user engagement and learning experience
* Developed a robust back end with NodeJS and ExpressJS, providing secure user authentication, course management, and seamless integration with Cloudinary for media handling
* Deployed the platform using Vercel for the front end and Render for the back end, ensuring scalability, security, and reliability of the application

## Achievements

- **Rank 3** in Hackathon AI-For-India Event, Build a Face Recognition Application using Python Certificate link
- Scored **97%** in Big Data Analytics: Opportunities, Challenges and the Future by Sebastian Binnewies, Lecturer, School of Information and Communication Technology, Griffith University.Certificate link
- Participated in Open Source Project Contribution **Hacktoberfest 2022**, making a total of **5 contributions** in webapp projects, improving the design and UI and adding features.
- Rank 1 in **hackerrnak bruteforce 2.0** Competition successfully captured the flag.
- Secured a position in the **top 4.2%** competitors in **Leetcode**, solving more than **300 problems in 100 days** in 2023 (handle: jayambe36).
- Rank 1 in the **ICT Ganpat University** on GeeksForGeeks (handle: jayambe36).
- Won the prestigious **COURSERA scholarship (financial aid)** in 2021, got a Machine Learning Course from Stanford University.
- Solved more than **1700+** chess puzzles on the official chess.com platform (handle: Chess).
- Deep dive into Space Science in terms of Math and Theory at **ISRO** - **Satellite based Navigation: A Journey from GPS to Mobile Phone Platform**.
- **Poster Presentation** During Hackathon at University **Student Grade Prediction using Decision Tree Algorithm**.
- Deep dive into Cyber Security, certified from International CyberSecurity Institute **ICSI** CNSS Certified Network Security Specialist at United Kingdom.

## Certificate

- **Spoken Tutorial (IIT Bombay) Certifications** Cpp Training, C Training, Advanced Cpp Training, BOSS Linux Training
- **IBM Developer Skills Network Certifications** Hadoop 101, Deep Learning Fundamentals,Machine Learning with Python, SQL and Relational Databases 101, DeepLearning.TV ML0115EN: Deep Learning Fundamentals
- **AWS Graduate Academy** AWS Machine Learning Foundation, Cloud Foundation
- **Cousera** Data Structures, Advanced Algorithms and Complexity, Machine Learning: Classification, Capstone: Retrieving, Processing, and Visualizing Data with Python, Introduction to Data Analytics for Business, Data Science Math Skills, Computer Vision Basics, Algorithmic Toolbox, Managing Big Data with MySQL, Mathematics for Machine Learning: Multivariate Calculus
- **EDX IBM** PY0101EN: Python Basics for Data Science, DS0105EN: Data Science Tools
- **Google Cloud** Google Cloud IAM and Networking for AWS Professionals, Managing Machine Learning Projects with Google Cloud
- **United State Institute of Peace** Introduction to Peace building Micro Course
- **Microsoft** Microsoft AI Classroom Series
- **Corporate Finance Institute** 3 Credit :Excel Crash Course - Spreadsheet Formulas for Finance