Apurva Mandalika

College Station, Texas | (979) 739-6155 | E-mail | LinkedIn | GitHub | Portfolio

SUMMARY

Data Scientist with nearly 5 years of experience delivering actionable insights through data analytics, stakeholder collaboration, and predictive modeling. Proficient in SQL, Python, Excel, and Tableau, with strong communication skills and a proven ability to translate business needs into impactful solutions. Experienced in building dashboards, performing statistical analysis, and supporting data-driven decisions across product and operations teams.

EDUCATION

Texas A&M University, *Master's in CS (CGPA 3.9/4)* *Aug 2023 – May 2025*
Amrita Vishwa Vidyapeetham, *B.Tech in CSE (CGPA 9.35/10), 8th rank in department* *Aug 2015 – May 2019*

SKILLS

SQL, Hive, MySQL, PostgreSQL, Python, Flask, FlashML, Data Cleaning, Data Wrangling, Data Mining, Machine Learning (Supervised and Unsupervised; Regression and Classification), Predictive Modeling, ETL, EDA, Feature Engineering (WOE, IV, etc.), pandas, NumPy, scikit-learn, PyTorch, Dashboarding, Tableau, Excel, Matplotlib, Seaborn, Plotly, Dash, Data Visualization, statistical analysis, A/B testing, hypothesis testing, Ad hoc Analysis and reporting, APIs, Java, Node.js, HTML, CSS, JavaScript, D3.js, Vue.js, Ruby on Rails, Swift, AWS(learning), Docker, Heroku, Git, GitHub, MS Office.

EXPERIENCE

Senior Data Scientist, [24]7.ai *Jun 2022 – Jul 2023*

- Introduced and implemented **Augmented Reality** - driven, video-based customer support solutions, improving customer problem **resolution rates** by nearly **50%** compared to traditional chat or voice-based support.
- Delivered multiple POCs and filed a **patent** for a novel feature, **published** in the **USPTO -** *PCT/IB2023/050635*.
- **Collaborated** with **cross-functional** teams to integrate solutions into the company.

Data Scientist, [24]7.ai *Jul 2020 – May 2022*

- Generated detailed ad hoc customer **analytics reports** in **Excel**, empowering stakeholders to make data-driven decisions.
- Proposed and built a comprehensive Model Performance Tracking dashboard using **Hive, SQL, Python and Excel**, that standardized evaluation processes for predictive models, resulting in a **40%** increase in actionable insights for client model performance assessments.
- Authored internal documentation using **JIRA - Confluence pages** to improve model transparency and onboarding efficiency.

Analytics Consultant, [24]7.ai *Jan 2019 – Jun 2020*

- Created Time On Page **(TOP) Prediction Model** and Page-Level Propensity to Purchase after Chat **(P2PC) Model**, increasing propensity to chat by **8%** and conversion rates by **12%**.
- Leveraged **SVM** and **Logistic Regression** in Python to create TOP models, optimizing customer engagement metrics.
- Conducted **data cleaning**, exploratory data analysis **(EDA)**, and **feature engineering** using Weight of Evidence **(WOE)** and Information Value **(IV)** for P2PC, ensuring robust predictive power.
- Used **Hadoop**, **Excel**, and **FlashML** to deploy scalable predictive targeting models, delivering actionable insights for diverse client use cases. Used supervised methods for **customer behavior modelling**; performed detailed EDA, **preprocessing**, and **anomaly filtering**.

PROJECTS

- **VitaFin: A Personal Health and Financial Data Visualization Dashboard (2025)** *Flask, Python, HTML/CSS, D3.js JavaScript.* Built a full-stack interactive dashboard to track and analyze personal financial and health metrics. Integrated real-time insights and benchmark comparisons, automating analysis delivery and visual storytelling for users.
- **Deep Learning Model for Image Classification (2024)** *Python, Jupyter Notebook, PyTorch.* Designed a hybrid deep learning model combining DenseNet and ResNet architectures for CIFAR-10 image classification, showcasing model architecture optimization, feature abstraction, and training automation. Achieved an accuracy of 92.5%.
- **Development and Comparison of ML and DL Models for Image Classification (2024)** *Python, Jupyter Notebook, PyTorch* Implemented Random Forest (44.97% accuracy), CNN (81.1%), and ResNet (83.6%) models to evaluate strengths and limitations on the CIFAR-10 dataset.
- **Multimodal Classification Model (2024)** *Python, Jupyter Notebook, PyTorch* Developed a fusion model combining a CNN for image data and an ANN for audio data to classify the multimodal MNIST dataset and achieved a validation accuracy of 98.92%. Emphasized representation learning and input heterogeneity.

ACHIEVEMENTS

- **Judge's Choice Award for 'Best Working Prototype'** at [24]7.ai's Global Hackathon
- Received the **'Team Excellence - Super Trooper' Award** at [24]7.ai's Global Annual Awards (2021).
- **Best Employee Awards** for Q3 FY21, Q4 FY22 & Q2 FY23.

CERTIFICATIONS

- **IBM Data Science Professional** on Coursera
- **Machine Learning** by **Stanford University** on Coursera
- **Introduction to Data Analysis Using Excel** by **Rice University** on Coursera