Prajwal Tidke

📞 619-520-9475 ✉ ms.ptidke@gmail.com in linkedin.com/in/ptidke9 ⭘ github.com/pTidke

EDUCATION

San Diego State University | *M.S in Data Science, San Diego, CA* | *GPA: 4.0/4.0* **Aug 2024 – May 2026**
Coursework: Machine Learning, Data Science, Data Engineering, Visualization

PDPM IIITDM Jabalpur | *B.Tech in Computer Science & Engineering, India* | *GPA: 3.2/4.0* **Aug 2016 – May 2020**
Coursework: Operating Systems, Data Structures and Algorithms, Object Oriented Programming, DBMS

WORK EXPERIENCE

Research Assistant, Part-Time **Feb 2025 – current**
San Diego State University San Diego, CA

- **Mental Health Chatbot for Construction Workers:**
  * Finetuning LLaMA using KBLaM (Knowledge-Based Language Alignment Modeling) with psychological knowledge triples to ground conversations in mental health contexts.
  * Developing a multilingual, privacy-aware chatbot tailored to construction workers, designed for deployment in low-resource mobile environments.
- **Core Temperature Prediction from Biosignals:**
  * Trained MLP, LSTM, and LagMLP models with temporal attention to predict core temperature from biosignals (PPG, ECG, EDA), achieving **25% MAE reduction** under standard evaluation.
  * Applied WGAN-based augmentation and Monte Carlo Dropout for uncertainty estimation; evaluated generalization using leave-one-subject-out cross-validation (LOSOCV).

Associate Data Engineer **Oct 2023 – Aug 2024**
TresVista Analytics Pune, India

- Managed a 2-person data analytics team for a major financial client, delivering reporting pipelines and automation solutions aligned with business objectives.
- Engineered robust ETL workflows using Databricks, PySpark, and Apache Airflow to consolidate and transform multi-source data, **cutting ingestion time by 50%**.
- Streamlined recurring financial model execution and Excel-based reporting via Python, VBA, and Power Automate Desktop, **reducing manual workload by 70%** and improving consistency.

Senior Data Engineer **Jul 2020 – Oct 2023**
LTIMindtree Pune, India

- Led a data engineering team of two as Project Lead, overseeing backlog planning, delivery governance, and cross-functional coordination for enterprise clients.
- Built and optimized ETL pipelines with SAP BODS to integrate datasets from SAP ECC, HANA, SQL Server, Oracle, and FTP sources into SAP BI warehouse, achieving **80% faster data delivery**.
- Automated manual data preparation processes using scheduled ETL jobs, improving operational efficiency and saving up to **1.5 FTEs** of effort.
- Managed SAP BODS environments, including job server setup, user roles, compliance, and patch maintenance under enterprise data governance standards.
- Led cloud modernization of legacy SAP pipelines by migrating to AWS Glue and architected Oracle-to-Snowflake transfers using SAP Workbench, with **50% projected cost savings**.

PROJECTS

Anomaly Detection in Tabular data | **Tech Stack**: PyTorch, Transformers (ALBERT), Autoencoders, WGAN-GP

- Developed a deep learning pipeline using WGAN for data augmentation and Deep learning models for anomaly detection and mutation pattern analysis in tabular biological datasets, with interpretability via attention-based visualization.

TrafficSensAI: Predictive Accident Analysis 🔗 | **Tech Stack**: DBSCAN, Random Forest, ArcGIS, React.js, Flask

- Built TrafficSensAI, a web app using React.js and Flask to predict traffic accident hotspots in San Diego via DBSCAN clustering and Random Forest classification, with ArcGIS-based geospatial visualizations.

TECHNICAL SKILLS

Languages: Python, Java, R, C/C++, Bash, VBA, Solidity
Machine Learning & Deep Learning: PyTorch, scikit-learn, Transformers, LLaMA, KBLaM, WGAN, LSTM, XGBoost, Random Forest
Cloud Platforms: AWS (S3, Glue, Redshift, EC2, Athena, DynamoDB, RDS, SQS, SNS, CloudWatch)
Data Engineering & ETL: Databricks (PySpark), Apache Airflow, SAP BODS, AWS Glue, Tidal Automation
Databases: PostgreSQL, MySQL, Oracle, SQL Server, Snowflake, MongoDB, Firebase
Visualization & Reporting: Tableau, Power BI, Qlik, SAP BI, ArcGIS
DevTools & Others: Git, Supabase, Notion, Confluence, Postman, ServiceNow, Microsoft Excel, DBeaver