TARUNKUMAR BANTUPALLI

Tempe, AZ • tarunbantupalli0601@gmail.com • (623) 281-4082 • linkedin.com/in/tarun-kumar06

EDUCATION

ARIZONA STATE UNIVERSITY   **TEMPE, AZ**
Masters of Science, Major in Data Science and Analytics   **DEC 2026**
(*Data Processing at Scale, Statistical Machine learning, Statistics for Data Analysts, Big Data*)   *CGPA* – 3.67/4

SRM UNIVERSITY   **Andhra Pradesh, IND**
Bachelor of Technology, Major in Computer Science; Specialized in Artificial Intelligence   **JUN 2024**
(*Python, Deep learning, DSA, Web development, Data Mining, Project budgeting*)   *CGPA* – 8.88/10

TECHNICAL SKILLS

- **Programming & Scripting:** Python, SQL, PySpark, Bash, Java
- **Data Engineering & ETL:** Airflow, AWS Glue, Lambda, Step Functions, Shell Scripting, Kafka
- **Cloud & Infrastructure:** AWS (S3, Redshift, RDS, EC2, Athena, DMS, EMR), Docker, Terraform
- **Databases:** PostgreSQL, MySQL, MongoDB, Snowflake
- **Analytics & Visualization:** RFM, PCA, DBSCAN, Regression, Tableau, Power BI, Matplotlib, Seaborn

EXPERIENCE

PROMPT ENGINEER   **HYDERABAD, INDIA**
ZENSAR TECHNOLOGIES PVT LTD –NVDIA PROJECT   **OCT 2024 - NOV 2024**

- Created and verified **400+ high-quality prompts** based on academic book images to support multimodal AI model training, ensuring **100% compliance** with linguistic, contextual, and formatting standards

DATA ENGINEER INTERN   **HYDERABAD, INDIA**
ELECTRONIC ARTS PVT LTD   **FEB 2024 - OCT 2024**

- Built and deployed **ETL pipelines** to ingest **10K+ daily records** from **Asset Panda API** into **MongoDB**, reducing processing time from **2 hours to under 10 seconds**
- Scheduled data pulls using **Task Scheduler** on **VMs**, enabling **100% automated refresh** of inventory data
- Designed **Power BI dashboards** via **ODBC connections** to MongoDB, delivering real-time visibility to **15+ business units** covering **2,000+ employees**
- Streamlined manual updates by developing a **Google Apps Script + JavaScript form** for inventory entry and validating API flows via **Postman**, with detailed logging for traceability
- Developed a **Slack bot** using **Flask**, integrated with internal systems to automate asset tracking, request submissions, and team-level queries
- Enabled self-serve workflows for business unit leads, reducing IT support requests by **60%** and eliminating dependency on email-based communications

PROJECTS

SCALABLE REDDIT ETL PIPELINE ON AWS

- Built a scalable ETL pipeline to process **1M+ Reddit posts/comments**, storing raw data in **S3** and transforming it via **Glue + Athena**, reducing processing time by **60%**
- Loaded structured data into **Redshift**, optimizing schema to support **sub-second queries** across **5+ analytical dimensions** (e.g., subreddit, sentiment)
- Automated data workflows with **10+ Airflow DAGs**, enabling **hourly refresh cycles** and fault-tolerant scheduling
- Implemented **metadata logging in PostgreSQL** for job recovery and validation, ensuring **zero data loss**

IOT SENSOR DATA PIPELINE FOR SMART ENERGY MONITORING

- Ingested **1M+ IoT energy sensor readings** from edge devices via **Kafka (or Kinesis)** streams into **S3** with real-time processing via **AWS Lambda**
- Aggregated time-series records into hourly/daily summaries using **PySpark**, and stored in **Amazon Redshift**
- Built **dashboards in Tableau** to visualize energy usage trends, anomaly detection, and peak load forecasting
- Achieved <5-minute data latency and enabled real-time alerting on spikes using rule-based triggers

CUSTOMER SEGMENTATION FOR ECOMMERCE MARKETING

- Clustered **541K+ customer records** using **K-Means, GMM, and DBSCAN** to identify marketing personas
- Enhanced segmentation by applying **RFM analysis**, boosting Silhouette score from **0.38 to 0.44**
- Used **PCA** for cluster visualization and selected optimal models via **Elbow method** and **internal metrics**
- Identified high-value cohorts for targeted campaigns, improving cluster cohesion by **30%+**

VOLUNTEER EXPERIENCE

STUDENT COUNCIL, SRM UNIVERSITY-AP   **SEP 2023 - NOV 2023**

- Developed a responsive, mobile-friendly event website using **HTML5, CSS3, JavaScript, PHP, and Bootstrap**, attracting **10,000+ visits** and **7,200+ registrations** in the first week