

NILANSHU TWINKLE

SENIOR SOFTWARE ENGINEER

ABOUT

Track-record of 6+ years in research & building end-to-end Big Data and AI/ ML Pipelines (large scale data ingestion (PySpark), data preparation (ETL), data analysis (EDA), feature engineering, technique selection & implementation, debugging, optimisation and deployment in production (CI/CD & Kubernetes)). Recognised functional capabilities in interfacing with global customers, rapid prototyping, R&D and meeting deliverables.

EDUCATION

- MS in Data Science Northwestern University Chicago US (2022 - 2024)
- B.Tech in ECE NIT Jalandhar Punjab (2013 - 2017)
- High School K.V. No 3 AFS Bathinda (2011 - 2012)

WORK EXPERIENCE

CGI

Big Data Engineer | Bangalore | Oct 2021 - Present

- CDL To Salesforce Data Migration: Designed and developed 15+ Azure Data Factory & Databricks pipeline (Spark RDD) migrating around 80 million+ (~2.5 TB) records daily for 5+ regions from Azure Data Lake Storage to Salesforce.
- Tire Claim Sync Process (Production): Developed and deployed middleware application with Python FastAPI for entire claim sync process and deployed on Tanzu Kubernetes and endpoints hosted on Google APIGEE and a daily job in ADB to sync claims to Salesforce tables and Salesforce Einstein for visualisation with Tableau CRM.
- Delta Table Sync: Daily sync of delta records from multiple on-premise MSSQL tables to Azure Data Lake after performing ETL with PySpark and then pushing to Salesforce.

WIPRO LTD

GNDC | Aug 2017 - 1st Oct 2021

- Ticket Transition Bot: Developed end-to-end pipeline for clustering (DBSCAN) and classification (LSTM) of ITSM records as business-critical tickets and EDA & Topic Modelling (NMF) for cognitive assist in focused resolution of 88% of critical tickets.
- Text Classifier: Achieved 85% accuracy in Service Now Ticket Classification using trainable BERT Embedding, XGBoost, Ensemble learning, GRU and deployed as Docker container on K8s in 5+ clients including USG Healthcare.
- Chatbot: Deployed NLU based Chat Bot and increased its usage from 8% to 32% in USG Healthcare by developing 6+ use cases like remote VM/ VDI restart, Cisco Device Profile Creation and Voice Mail Setup, ServiceNow ticket expedition & status check. Used Microsoft Bot Framework, Python and Rivescript for designing complex chat workflows and Spacy NER for Intent identification, entity extraction, text analytics.
- PII Detection BOT: Designed and developed 40+ multi-region Personally Identifiable Information (PHI, PCI, PII) detection bot with Regex and Spacy and Spark based NER (Boltzman machine) incremental predictive modelling.
- Cognitive Search Bot: Part of team in developing Elasticsearch based bot to fast index and search 10000+ Knowledge Articles and millions of ITSM tickets in near real time.

PERSONAL PROJECTS

- Automated Freeway Deduction Service: Designed and developed an Python FastAPI based application for Freeway Toll Deduction using Automatic Number Plate Detection and RFID data with velocity of 430K request/day for single expressway with deduplication of multiple requests (with Redis) and near real time settlement with Banks using decoupling (using RabbitMQ) with and simultaneous checks of Violations, exemptions, blacklisting. Missed entry and exits are also handled and all of it deployed on Kubernetes with Event driven autoscaling. Logging with Elastic-Logstash-Kibana and monitoring with Prometheus
- Topcoder - EdgeNet Shoe Classification: Developed shoe image classification using Convolution Neural Net (CNN) and Transfer Learning with ResNet (V2) model (Keras) to identify different types of shoes with 83% of accuracy.
- Topcoder - NASA QGroundControl System - Human Factor Client Implementation: Mimicking user interactions with the QGroundControl UI platform based on given dataset having information about user interactions with different IO devices and gaze data.



CONTACT

P: +91-8699498680
E: nilanshu01@outlook.com
L: linkedin.com/in/nilanshu-twinkle-72a1a1126
C : credly.com/users/nilanshu-twinkle

PROFESSIONAL SKILLS

- Deep Learning (Tensorflow/ Keras), NLP, Markov Chains
- RNN (LSTM/ GRU), Attention Framework, Transfer Learning, Transformers, CNN
- Clustering, Topic Modelling, SVM, Decision Forests, KNN, Dimensionality Reduction, PCA
- Pandas, Numpy, Spacy, Scipy, Matplotlib, Seaborn
- Py-Spark, Hadoop, Data Lake
- EDA with R, Python & Tableau
- Azure Databricks & Data Factory, AWS Sagemaker, AWS Textract, Athena
- Docker, Kubernetes, Istiod
- SQL, PostgreSQL, NoSQL
- Elasticsearch-Logstash-Kibana, Redis, RabbitMQ
- REST(FastAPI), SOAP, APIGEE
- GIT Pipelines, Azure Devops
- Terraform, Ansible

MICROSOFT AZURE CERTIFICATIONS

- Data Scientist Associate (DP-100)
- Data Engineer Associate (DP-200, DP-201, DP-203)
- AI Engineer Associate (AI-100)
- Azure DevOps Expert (AZ-400)
- Solution Architect Expert (AZ-303, AZ-304)
- Developer Associate (AZ-204)