Shweta

<http://www.linkedin.com/in/shwetasinghchandel>

# EDUCATION:

**Stevens Institute of Technology, Hoboken, NJ |** Master of Science in Information Systems (data science) | **Aug 2019 – May 2021 Uttar Pradesh Technical University, India |** Bachelor of Technology in Electronics and computer Engineering| **Sep 2009 – May 2013**

# SOFTWARE SKILLS

**Languages**: Python, R, SAS, SQL,Java, JavaScript, Scala, C, XML, Html5.

**Python Libraries:** Keras, TensorFlow, Numpy, Pandas, NTLK, SciPy, PyTorch, Pyro, OpenCV, Matplotlib, Geopandas, Spacy

**Frameworks, Libs, Version Control:** Nodejs, GraphQL, Presto Query Engine, Git, GitHub, Bitbucket,MLflow

**Databases & Tools:** MongoDB, PostgreSQL, Hadoop, Airflow, Spark, Hive, Snowflake, Salesforce, Tableau, Quicksight, CUDA

**Cloud Technologies:** GCP (Pubsub, GKE, Compute Engine, Spanner, BigQuery, Composer), MS Azure, AWS (S3, IAM, EC2, DynamoDB) **AI/ML:** Regression, Classification, Clustering, Feature Engineering, Neural Nets, Deep Learning (CNN, RNN, LSTM), Natural Language Processing (BERT, NER), Decision Tree, Naïve Bayes, KNN, LLM, Causal inference, RAG, Transformers, CUDA

# PROFESSIONAL EXPERIENCE:

## VISYS CLOUD TECHNOLOGY, California | Role: Sr Data scientist/ML Engineer May 2023 - CURRENT

* Collaborated to Build GenAI chatbot using GPT and Google Vertex AI, utilized RAG (Langchain, Haystack) with prompt engineering.
* Fine-tuned BERT LLM (Hugging face transformer) for NLP tasks as Named Entity Recognition (NER) to find key terms in documents.
* Designed solutions for error handling, ensuring the robustness of NLP-based applications in real-world use cases.
* Developed a mail-classifier (text classification) for support team using TF-IDF, word-to-vector models, with 86% accuracy rate.
* Designed and implemented complex data models to support business intelligence, reporting, and analytics applications.
* Performed hypothesis A/B testing to study impact of promotions across merchant channels like salesforce for cost optimization.
* Deployed, tested, monitored ML model on GCP using MLOps frameworks Docker, Kubernetes, Airflow, Kubeflow, Argo Workflows.

## FACEBOOK (contract via EPITEC), California | Role: Business Data Analyst 2 March 2022 - March 2023

* Built Prediction model using python (regression, SVM) to forecast customer behavior and provided stakeholders insights using Tableau, Increased customer conversion rate by 21% by recommending improvement to supply chain enterprise team of Facebook.
* Developed Performance matrix using SQL, Python, Tableau to monitor KPIs, maintained SLA, Performed UATs, for recommendation.
* Developed Python scripts for AI data validation, using pandas, optimized SQL queries and built ad-hoc data pipeline and reports.
* Analyzed performance of new hardware using Python & Presto SQL to identify & bucket requests into categories, for feature launch.
* Collaborate with stakeholders to gather requirements, design, develop, maintain user-friendly & scalable systems, ML applications

## BRIGHT MIND ENRICHMENT, New York | Role: Data Science Engineer September 2021 - March 2022.

* Leveraged Geospatial Data, Random Forest, Regression and K mean clustering to convert maltreatment risk predictions into actionable intelligence that stakeholders can use to efficiently allocate limited child welfare resources
* Built sentiment classification model using NLP techniques on reviews such as word embeddings (word2vec, TF-IDF), vector search.
* Developed a deep learning image classification system for aid categorization, reducing manual sorting time by 30%. trained a CNN using TensorFlow and Keras, achieved 90% accuracy by transfer learning (ResNet50, InceptionV3), data augmentation, and hyperparameter tuning and learning rate adjustments to optimize model performance.
* Deployed, monitored ML applications (CI/CD) on AWS utilizing PyTorch, TensorFlow and MLOps deep learning frameworks (Caffe).

## GRAVITE EDUVENTURE, India | Role: Decision Scientist February 2018 - June 2019

* Implemented a time-series forecasting model using Prophet and BigQuery ML to predict demand trends, 12% reduction in stockouts.
* Built a recommendation system using XGBoost in Python, improved accuracy by 25% through advanced feature engineering (PCA).
* Developed a customer segmentation model using clustering algorithms (K-means, DBSCAN) to identify customer behavior groups.
* Designed A/B test with chi-square, Bayesian analysis to assess marketing strategies, optimized Ad targeting with correlation analysis.
* Developed ETL (Extract, Transform, Load) pipeline to scrape data using Pyspark REST APIs in JSON format, loaded them in MongoDB.

## HCL TECHNOLOGIES, India| Role: Analyst July 2016 - January 2018

* Optimized Advanced SQL query by indexing and normalization to improve data retrieval, reduced processing time by 25%.
* Developed ETL, ingesting ~5Tb/day from iceberg, hive using PySpark, performed Spark optimizations to reduce runtime by 66%.
* Developed a customer churn model (logistic regression, SVM, Decision tree, Naïve bayes) using python, improved accuracy by 25%.
* Conducted fraud detection using anomaly detection algorithms in R and SAS, identifying fraud transactions with 90% precision.
* Performed data wrangling on large data sets using python, SQL and monitored KPI with Power BI to provide business insights.
* Performed ANOVA, t-test, chi-square test, correlation analysis using Excel (Pivot table, index, array), VBA, to generate data insights.
* Designed Star, Snowflake Schema Data Models for Data Warehouse, created data dictionaries and maintained documents

## J K T CONSULTING Pvt Ltd, India| Role: Software Engineer June 2015 - July 2016

* Designed ETL workflows in SSIS and Talend to integrate data from different data sources, MySQL, SQL Server, PostgreSQL.
* Executed SQL stored procedure, designed data profiling packages in SSIS to uncover data quality issues, Developed ETL test scripts.