

Chetan Chauhan

chetanknmiet@gmail.com | +1 469-739-0777 | 6656 Plaza Via 320B, Irving Tx, 75039 US

Having L2S Visa valid till 31-Aug-2025

EDUCATION AND HONORS

National Institute of Technology, Surathkal

Master of Technology, Computer Science (7.58/10 CGPA)

Karnataka, India

July-2014

Uttar Pradesh Technical University

Bachelor of Technology, Computer Science (76.74%)

Uttar Pradesh, India Honors

May-2011

SKILLS

- **Programming Languages and Framework:** Python, Tensorflow, Bash, Shell Script.
- **Operating System:** Windows, Linux/Unix.
- **Big Data Technologies:** Spark, Pyspark, HDFS, Kafka, Apache Airflow, SparkNLP, Databricks.
- **Cloud Technologies:** AWS S3, EC2, Lambda Function, AWS Sage Maker, AWS Glue, Azure Blob Storage, Azure Service bus, Azure Event hub, Azure Databricks, Azure Functions, Azure Data Factory, Open Stack Cloud.
- **Databases:** MySQL, MongoDB, Elasticsearch, Redis, Neo4j.
- **Machine Learning Packages:** Scikit-learn, Pandas, Numpy, NLTK, Spacy 2/3+, Fasttext, TextBlob, Stanford CoreNLP, Gensim, Stanford Stanza, SciSpacy, Apache-Ctates, RASA, RASA NLU, RASA Core, Tesseract, FBProphet, HuggingFace, Snorkel Flow.
- **Machine Learning Models:** Decision tree, Random Forest, XgBoost, Logistic regression, Linear and Non-linear regression, SVM, Neural network, Clustering, Glove, Word2vec, FastText, Multilingual Word Embedding, Bert Embedding, Knowledge Embedding (TransE, TransH, TransR, TransX), OpenKE, OCR, Hypothesis Testing, A/B testing, Probability distribution, Random Process, Healthcare Ontology (UMLS, RxNorm, LOINC, SNOMEDCT etc)
- **Visualization Libraries and Tool:** Matplotlib, Seaborn, Apache Superset, Kibana, plotly.
- **Other Miscellaneous ML Technologies:** Text Annotator, LOBE, Label Studio, Yedda Annotator, Doccano annotator, Microsoft Vott Image annotator, Pycharm, GIT, Bitbucket, JIRA, PDB, Docker, MLFlow, CML.
- **Area of Interest:** Information Retrieval, Natural Language Processing, Recommendation System, Algorithm Trading, Chatbot, Machine learning, Deeplearning, Augmented Reality, Data Structure and Algorithm.

PROFESSIONAL EXPERIENCE

Evolent Health International

Lead NLP Data Scientist, Risk Analytics Team

Pune, India

Aug-2020 – Dec-2022

- Mentoring and leading a team of four Data Scientist.
- Owned end to end responsibilities for deploying the models into production on Azure cloud.
- Responsible for all the model's lifecycle management, HIPPA compliance, risk assessment etc.
- Developed Suspecting predictive engine and created FastAPI rest endpoint for downstream products.
- Created an augmented intelligence based medical **chart review accelerator tool** which helps in increasing the productivity of medical coders **by 3X**.
- Increased the reimbursement amount received by payers from **CMS by 10%** while identifying missed diagnosis codes present in medical charts with the help of chart accelerator NLP tool.
- Developed pipeline for Healthcare Contract extraction from large volume of PDF contracts. **Save 8000+ hours** with this AI automation tool.
- Developed Airflow Dags to process Prior Authorization forms and Claims form using AWS textract, OpenCV and Tesseract.

Cognogix Technology (Client: Icertis) Contract

Lead Data Science Engineer

Pune, India

Apr-2020 – July-2020

- Developed a **custom NER for legal entities** recognition for contract management platform.
- Developed a contextual model to find the similar clause present in the contract with the clause libraries.

UBS Group AG

Data Scientist, Insight and Innovation Team

Pune, India

Jan 2018 – March 2020

- Developed a **Service Now Incident ticket triage system** which helps to identify the right assignment group. 80% of the self service ticket were being triaged through this system.
- Redesigned the search algorithm in the **UBS portal knowledge article search engine**. Developed an algorithm to find the relevant knowledge article based on user query.
- Decreased the bounce rate to 20% and increased the number of clicks to 30% after deploying search algorithm.
- Developed an **Organizational Network Analysis** graph using NEO4J on employees outlook meeting data and distributed group list data.
- Created an analysis report for employee's collaboration with the team and outside the team using graph network analysis metrics.
- Developed **Hot Roles jobs analysis dashboard** for HRBP with the help of crawled Job descriptions from various job portals.
- Developed an automated **HR FAQ chatbot** which can engage, empower and delight your employees, and liberate your HR team from routine and mundane enquiries
- Developed a **Commentsopedia survey analysis tool**. This tool analyzed the employee survey free text data and allow to concentrate on important issues being raised within various sector and business unit.

DB Corp

Sales and Marketing

Noida, India *Data Scientist*,

Sep-2016 – Dec- 2019

- Developed a **facebook news reach** predication model which helps in deciding which news gets viral in near future.
- Developed **news recommendation engine** for both mobile app and news website.
- Created big data based recommendation algorithm using both content and user reading pattern. Increased the user engagement time to 30%, decreased bounce rate, increased the CTR.
- Created a personalization recommendation algorithm on DB Corp mobile app using novel algorithm based on association rule mining and page rank.
- Developed and created A/B testing experiment in google analytics to compare recommendation algorithm performance or to compare any design changes.
- Developed Ad CTR prediction model for finding the CTR of Ads running on various DB Corp news portal.
- Developed a big data pipeline using apache-spark for handling large amount of user click data.
- Analyzed user behaviour, customer segmentation with cohort analysis, market basket analysis and RFM analysis.

Innoplexus Consulting Pvt Ltd.

Gurgaon, India *Data Engineer*

March-2016 – July 2016

- Created an **advanced text search** for one of the inhouse healthcare product.
- Developed a model to extract adverse drug events from clinical research raw text data.
- Developed an aspect-based sentiment analysis model to find the sentiment of drug adverse reaction on social media.

One Convergence Pvt Ltd.

Software Engineer

Hyderabad, India

July-2014 – Feb-2016

- Worked on High performance computation hardware to analyze the streaming data coming from sensors present on FPGA card.
- Developed a real time anomaly detection algorithm to triggered an alert based on sensor data.
- Created a **license manager for F5 Load balancer** as service in Open stack cloud.
- Created streaming data dashboard with opensource tools like collectd, graphite, kibana, elasticsearch.

Master Thesis

Teaching Assistant

Surathkal, India

July-2013 – May 2013

- Automated Behavioral based malware detection using machine learning.

Certifications

1) *Neural Networks and Deep learning (7NAZ3DW4N9TZ)* 2) *Ontology for business Analysis*

Achievements and Extra curricular activities

- 1) GATE qualified 99.44 Percentile.
- 2) Rising star awards in One Convergence, DB Corp, UBS, Evolent.