

# Abhishek Jain

Sr. Data Scientist, Gen-AI Enthusiast

Pune | +91-9545058755 | [abhishekjainin88@gmail.com](mailto:abhishekjainin88@gmail.com)

## Summary

- 11+ years of experience with more than 8+ years of experience in Machine Learning, NLP Predictive Modeling, Design, Development, Testing, & Deployment along with 4 years' experience in Sql Development.
- 8+ years of experience working on agile methodologies with experience and knowledge of project management, SDLC methodologies including Waterfall and Agile
- Experience in preparing the functional and Business specification based on the requirements.
- Ability to influence stakeholders and work closely with them to determine acceptable solutions.
- Strong understanding of Data Analysis, Data Modeling, Data Ingestion, Data Reporting and ML/AI Pipeline.
- Manage effective communication with various stakeholders.

## Technical SKILLS

**Machine Learning** Regression, Classification, clustering(K-means), Principle Component Analysis

**Natural Language Processing** TFIDF, word2Vec, Spacy, NLTK, T5, BERT, Transformers, & LLM

**Deep Learning** ANN, RNN, LSTM, Bi-LSTM, Transformer

**Statistics** Hypothesis testing, T test, Z test, ANOVA (F test), Chi-square test

**Ensemble Methods** Bagging, Boosting

**Programming Language** Python, R, C#, C, SQL

**ML Framework** Tensorflow, Keras, Pytorch, Sklearn, Hugging Face

**AI Cloud Platform** AZURE ML, AWS-EC2

**Database** Oracle, SQL Server, Mongo DB, Cassandra

**Deployment** Flask, Docker & Kubernetes, Server less deployment (Aws Lambda, Azure function)

**Domain** Banking and Finance, Telecom, Supply Chain, HealthCare, Legal and Banking

**Software Tools** Jupyter, Git, Spyder, Excel, R studio, Putty, Winscp, Postman, Visual Studio, PyCharm, Visual Studio

**Operating System** Windows 10, Ubuntu 18

## Work experience

**Northern Trust** (Second Vice President)

Aug-2020 – Present

### To predict the clients who may leave NT

As a Financial Institution Northern Trust needs to increase their clients and at the same time the maintain the relation with existing clients so that they don't lose business. Goal is to predict the clients which may leave in next year. This will help NT to take necessary action to retain them as clients. (Python, Scikit-learn, Attrition Model using XGBoost, Random Forest and Clustering)

### Dissection of Document for Due Diligence

NT also works as a Private Equity firm. As a PE firm it has to select a firm in which NT can invest, operate and increase revenue to sell them with increased valuation. This need due diligence to be done at each level. Goal is to dissect a bunch of document into different section which can be forwarded to specialist in each field to perform due diligence at their level. (Python, Scikit-learn, NLP, BERT)

### Contract Management Tool (CM-Tool)

Created a web service to extract various legal clauses from contract documents. Developed 5+ NLP models to perform task like Text Extraction, Summarization, Generation and Classification using algorithm like TFIDF, word2vec, T5, & BERT used

framework like TensorFlow, PyTorch & deployed on AWS cloud by doing performance optimization and scaling with Nginx, Gunicorn, SQS and Docker.

## Conversational Chatbot

This chatbot utility facilitates document upload for querying purposes. Users can ask document-related questions, and the chatbot, will provide relevant responses based on the document's content. The libraries used are deep learning LLM, LangChain, Python and vector databases Chroma.

## Cognizant Technology Solution (Senior Associate)

Apr-2016 – Aug-2020

### Wafer Fault Detection

A telecom organization has a production line set up with multiple wafers. If there is any issue in the production line, then is very difficult to find out the faulty wafer as they have to stop the production line to check the faulty wafer.

Goal is to detect which wafer is faulty so only that section of line will be closed and particular wafer is replaced.

### PO Report Entity Extraction

The project involves extracting various entities, including Vendor Name, Date, Product ID, Amount, etc. from the PO Report. The extracted information is then stored in Excel and circulated to the relevant team through email. Additionally, tasks include converting image PDFs to searchable PDFs using OCR, building a NER extraction model for entity identification, and deploying a web service using Flask on the server for seamless functionality.

### Sentiment analytics Solution to predict Sentiments

Sentiment analytics solution on social media data for Norton Products. Roll up the SFDC case comment data to customer level and predict the sentiments using TFIDF, Random Forest. It will help us to derive the promoter & Detractor's customers.

## Accenture Services Pvt. Ltd. (Senior Software Engineer)

Sep-2010 – Feb-2016

### Replace written prescription of a patient with electronic document

Replace written prescription of a patient with electronic document which can be used among different faculties (doctors, physicians, nurses, pharmacist etc.) based on different workflow rules.

### Sql Developer

Analysis, requirement gathering and development of modules, Development, Testing, Deployment, attend daily calls with Onshore to discuss the logs and defects, mentoring junior developers, Reviews (peer and group), test data creation, retrospective and Releasing changes to UAT and Live.

## Education

**Sagar Institute of Research and Technology, Bhopal** Bachelor of Engineering in Computer Science - 73.03%

**SBKV Govt. Sarvodaya Vidhyalay (CBSE)** 12th - 73.0%

**D.A.V. Centenary Public School (CBSE)** 10th - 71.4% from

## Personal Details

Date of Birth: 7 Mar, 1988

Father's Name: Devendra Jain

Languages Known: English, Hindi

Permanent Address: BS-2, Divya Heights Rahatani, Pune - 411017