

ABHISEK DAS

+(91) 7278650267 ♦ Kolkata, West Bengal, India

[E-mail](#) ♦ [LinkedIn](#) ♦

EDUCATION

Bachelor of Technology (Computer Science), West Bengal University of Technology

2010 - 2014

SKILLS

AI Skills	Machine Learning, Deep Learning, NLP, LLM, Time Series Analytics, EDA
Languages	Python, C, C++, Java
Libraries	Numpy, Scipy, Pandas, Matplotlib, Scikit-learn, Seaborn, Keras, Tensorflow, LangChain
Cloud Technologies	AWS

EXPERIENCE

Data Scientist

Oct 2019 - Present

Tata Consultancy Services (Innovation Labs)

Kolkata, India

- **Unsupervised Remaining Useful Life (RUL) Prediction of engines**
 - Training data is divided into healthy and non-healthy parts and used to train an **LSTM Encoder-Decoder** model for a machine's life-cycle.
 - The model is then applied to the full dataset, generating high error values for unhealthy parts and low error values for healthy parts. These errors are reverse-normalized to assign a Health-Index (HI) between 0 and 1. Comparing the error of a new test instance with previous HI values estimates the Remaining Useful Life (RUL), leading to **25 % reduction in operational expenses**.
- **Estimation & Causal Analysis of Engine Failure of leading US Manufacturer**
 - Used **Fuzzy based approach** for augmenting the existing dataset
 - Applied Deep Learning and machine learning techniques such as **LSTM** and **RFC** to estimate actual failure dates of engines with **~ 80 %** accuracy
 - Used **Explainable AI (XAI)** library such as Lime on the datasets to find out leading causes of failure.
 - Collaborated with on-site client team members to collect, process and present the data to all the stakeholders. Received appreciation from CIO of client.

Jr. Data Scientist

Sept 2015 - Oct 2019

Tata Consultancy Services (Innovation Labs)

Kolkata, India

- **Unsupervised Clustering of sensors based on time series data similarity for leading hardware manufacturing giant**
 - Delivered a working PoC solution for unsupervised clustering of **700+** sensors based on time series data to remove redundancy.
 - Applied **Pattern Recognition** techniques like **SAX** is used to encode the time series data.
 - Dynamic Programming based **Longest Common Subsequence (LCS)** approach is then used based on some threshold to identify the similar sensors and put them into respective clusters.
- **Unsupervised identification of loops in flight paths of aerial vehicle engine by UK manufacturing giant**
 - Loops in flight paths create operational overheads for airline operators. Thus, a need to identify loops in flight paths was required by customer to increase fuel efficiency.
 - Used **SAX encoding** on derived features to encode the input data consisting of **5000+** source-destination pair of airports. Used **Google Maps API** for validating the output of our approach and presenting our results. Our analysis helped reduce client's operational expenditures by **30 %** . Received appreciation from all stakeholders of the project.
- **Building analytics capability to existing IoT platform.**
 - Contributed towards writing wrapper classes for different analytics algorithm like **SVM**, **RFC**, **LSTM Encoder-Decoder** , etc. and packaged it into libraries.
 - Built unit test cases with sample codes and documentation of the libraries. Followed best practices while writing the codes and documentation.