

DEEPAK SINGH

[deepsai8.github.io](https://github.com/deepsai8) | [linkedin.com/in/deepak Singh8819](https://www.linkedin.com/in/deepak Singh8819) | +1 (628) 468-7621 | deepak.id1988@gmail.com

PROFESSIONAL EXPERIENCE

Data Scientist **Boston Children's Hospital** **San Francisco, Nov-2022 to Present**

- **Abstract accepted at American Clinical Neurophysiology Society (ACNS):**
 - Nonlinear Dynamics and Tensor Factorization for EEG (ElectroEncephaloGraphy) Biomarker Discovery as a Data Science Intern under the supervision of Dr. William Bosl
- Implemented Tensor Factorization in python for multimodal data fusion to predict presence of BECTS (Benign Epilepsy with CentroTemporal Spikes) using EEG data from 26 each of control and test subjects aged 2 to 19 years, with a 0.92/0.11 AUROC/Brier score, with awake segments and 0.92/0.11 AUROC/Brier score, under sleep stage N2

Data Scientist **Accenture** **India, Apr-2018 to Jun-2022**

- **Billing Classification model:** Spearheaded design and development of ML based classification model, which **detects erroneous/non-billable claims**, which improved working capital by 10% (~ \$ 1.6 M) y-o-y in procure-to-pay process of supply chain
- **Predictive Performance Model:** Deployed a performance prediction model with a MAPE under 5% for over 40k partners of a billion dollar client, which enabled B2B partners with more selling time than managing compensation
- Developed a **payment optimizer** model which **recommends key-focus customers**, basis historical data, resulting in 15% improvement in average payment terms in order to cash (OTC) cycle
- **Data-driven contract management:** **Designed an SQL backed** low-cost online workflow system that converted unstructured contract data into meaningful insights, and transformed contract experience for 5000+ customers,
- Enhanced **online catalog through experimental design as a part of omni-channel strategy** for a client, which helped drive demand for products & services by 3x and provided a seamless experience for its B2B customers
- Lead **cross-functional technical & business teams to operationalize the automated onboarding process** of ~30k business partners, enhancing productivity by 45% and giving a 99.5% onboarding accuracy

Senior Engineer **Hindustan Petroleum** **India, Jun-2011 to Apr-2017**

- Designed and developed a **predictive cost estimation model** for parametric cost estimation to enhance estimation accuracy by more than 80% and thus saved \$100k per annum
- Built and implemented an **automated invoice validation system** to reduce the validation time from 3 days to less than an hour per invoice, and thereby achieved savings worth \$70k annually
- Modernized document search across multiple functions, by deploying OCR and **key-word classifier on all** documents to reduce search time by 30%

EDUCATION

University of San Francisco **MS in Data Science** **Jun-2023, San Francisco**

Courses: Machine Learning, Deep Learning, Linear Regression, PCA, Time Series Analysis, A/B Testing, Visualization, Distributed Computing, Relational Database Management Systems, NoSQL, Data Structures and Algorithms

Indian Institute of Management Calcutta **MBA** **Mar-2018**

Maharshi Dayanand University **Bachelor of Technology** **May-2009**

SKILLS

Python (scikit-learn, numpy, pandas, scipy, matplotlib, seaborn, jupyter), SQL (PostgreSQL, BigQuery), NoSQL (MongoDB), nltk, spacy, A/B Testing, Linear Regression, Time Series (ARIMA, Prophet, VAR), LSTM, Web Scraping, Distributed Computing (Apache Spark, Apache Airflow, pyspark), Cloud Computing (GCP, Databricks), MLFlow, PowerBI, Power Apps

PROJECTS AND ADDITIONAL

- **Market SentiMeter:** Developed a market sentiment aggregation pipeline, where users can assess any stock and calculate the propensity to buy/hold or sell based on Financial and Technical data, social media sentiment and corporate announcements by the organization
 - Apache Spark, Airflow, MongoDB, pymongo, plotly, SQL (postgres, psycopg) Databricks, MLFlow, django
- Developed a news insights visualization through New York Times API, captured news article data from 1991 onwards, analyzed news article trends from various locations, topics and organizations, using Apache Spark, Google Cloud Platform, Databricks and Plotly
- Enthusiast 3D model artist & Game Developer with online presence at [Sketchfab](#) (11,000+ views) and [Itch.io](#)