

Arnab Samanta

ACADEMIC PROFILE

DEGREE	INSTITUTION	CGPA	YEAR
IDD (B.Tech + M.tech) Metallurgical Engineering	IIT BHU	8.69/10	2019-2024
CBSE 12th Boards (School Topper)	Auckland Public School	94.6%	2019
CBSE 10th Boards (School Topper)	Army Public School, Mamun	10/10	2017

Languages	Python (Numpy, Pandas, Scikit-learn), SQL
Technologies	Tensorflow, PyTorch, Selenium (Web Scrapping), Seaborn
Data Skills	Data Analysis, Statistics, Data visualization, Machine learning, Deep Learning

PROFESSIONAL EXPERIENCE

Trademo DS Intern (May'23-Aug'23)	<i>Worked on SaaS products for a new vertical in a Supply Chain Intelligence Startup</i>
	<ul style="list-style-type: none">Automated an extreme classification task (Trademo Classifier) for the global customs offices & e-commerce.Designed a 12-language multilingual translator to help establish a Vector Database for HS Code classifier.Leveraged FastAPI, Torch, MLFlow with Information Retrieval, Product Quantization, Siamese NN, HNSW using Data Flywheel concept integrating user prompts & product category to craft a production level classifier.Devised a robust HS code misdeclaration classifier (Trademo Watch), refined resource use via data-driven metrics.Achieved an 81% accuracy for top 3 HS Code recommendations and achieved a 83% ROUGE score on the translator.
Nuartin Labs ML Intern (Jun'22-Jul'22)	<i>Product Digitization for \$413B Retail Chain Supply</i>
	<ul style="list-style-type: none">Operated on multilabel text classification problem of product description derived from e-commerce product images.Performed text cleaning, text augmentation & created ensemble models achieving a macro recall score of 0.9613.

PROJECTS

ML Case Study Oct'22 - Nov'22	<i>Delivery Partner Deployment Optimisation: Swiggy</i>
	<ul style="list-style-type: none">Decreased rejection rates of orders by aligning them with the right set of delivery partners at the ideal moment.Conducted EDA on Swiggy's data, unveiling rejection trends concerning distance-time dynamics and zone effects.Employed data preprocessing indulging feature engineering, imputation, curation and extracted important features.Attained a precision-score of 0.93 with a Random Forest Classifier and devised an efficient real-time solution.
Self-Project Dec'21 - Jan'22	<i>Music Genre Classification</i>
	<ul style="list-style-type: none">Built a genre classification model on GTZAN dataset consisting of 1000 songs of 30 secs each from 10 genres.Converted audio samples from stereo to mono channel, extracted 13-dim MFCCs from the segmented audio files using librosa and implemented an ANN model, CNN model and stacked LSTM model with Adam optimizer.Achieved best test accuracy of 73.97% with CNN model trained for 100 epochs and a learning rate of 0.0001.
Self-Project Mar'22 - Apr'22	<i>Multi-Way Classification of Semantic Relations</i>
	<ul style="list-style-type: none">Implemented the SemEval 2010 Task-8 using the annotated sentences dataset having a total of 10 classes.Preprocessed text, used spacy's embedding model to create vectors of tokens and padded each sentence with 0s.Built a PyTorch model with a Bi-directional LSTM of hidden size 128, a multi-head attention layer, a linear layer consisting of positional information of entity pairs and a softmax layer trained with Adam optimizer for 100 epochs.
Self-Project Sep'21 - Oct'21	<i>Human Resource Analytics</i>
	<ul style="list-style-type: none">Analyzed a MNC's dataset to predict a potent employee's promotion to save time in the transition period.Applied data transformations, used SMOTE to oversample and selected relevant features using correlation matrix.Selected XGBClassifier from 6 other models with the help of cross validation, tuned 6 parameters of XGBClassifier using optuna for 150 epochs and achieved an accuracy of 94.380% on validation set.

POSITIONS OF RESPONSIBILITY

Varanasi WordFest Apr'21-May'21	<ul style="list-style-type: none">Tab Director, IIT BHU PD 2021<ul style="list-style-type: none">Maintained the digital logs for a 72 hour live international tournament with 50+ teams across Asia.Saved 5+ hours by automating data entry of 250+ participants with Python script and deployed on Heroku.Worked with a team of 5 and coordinated with the event judges to ensure no delays in the tournament.
Literary Club, IIT BHU Aug'19-Present	<ul style="list-style-type: none">Mentor at Literary Club, IIT BHU<ul style="list-style-type: none">Held numerous competitive debating and adjudication workshops with more than 100 participants.Organised practice debating tournaments for students and guided them in inter-college competitions.Managed and taught technical requirements to host large-scale international debating tournaments.

HONOURS AND ACHIEVEMENTS

Technical Achievements	<ul style="list-style-type: none">Winner of Cassandra held under UDYAM'22 on kaggle, a Data Science event participated by 60+ teams across India.Led a 10-member team in Digital Alpha's SEC Filing Analyzer, a high prep event at annual Inter IIT Tech Meet 10.0.Ranked 3rd in Goldberg's alley held under the extreme engineering team events at Technex 2020 IIT BHU.Ranked 25th out of 2132 (top 2%) in Prediction of genetic disorders hackathon conducted by hackerearth.
Extra-curricular Achievements	<ul style="list-style-type: none">Bagged 1st position in group singing event at Aagman 2019, the inter-branch fresher's competition.Winner of Hindi story writing competition conducted under Aaftaab 2020, annual literary fest of IIT Goa.