PRABAL SABARWAL

ACADEMIC QUALIFICATIONS

Year	Degree /Board	University /Institution	%/CGPA
2025*	Post Graduate Diploma in Business Analytics	IIM Calcutta, IIT Kharagpur, ISI Kolkata	-
2015	Dual Degree - B.Tech + M.Tech. Chemical Engineering	Indian Institute of Technology, Madras	7.69/10
2010	CLASS XII	Stepping Stones School, Chandigarh	82.4 %
2008	CLASS X	Stepping Stones School, Chandigarh	89.3 %

KEY SKILLS/TOOLS	Data Science, Statistical Data Analysis, Tableau, Excel, SQL, R Programming, Python, PyTorch, Scikit-Learn

ACADEMIC PROJECTS

Airbnb Price Prediction (Regression)
- Predicted **listing price** for **75k+** Airbnb properties (**26** features) with **linear** and **polyno.** regression models
- Treated **outliers**; Tackled **multicollinearity** (**VIF** and **backward selection**); Applied **Ridge regularization**
- Analyzed residuals with **KS test** & **QQ** plot | Built **DT, RF** & **XGB** regressor; ↑R^2 by **23%** to **0.75** with XG Boost

Crude Oil Price Prediction Model (Time Series)
- Forecasted crude oil prices using **28** years of WTI data (**300+** obs.); analyzed trend & **stationarity** with **ADF**
- Used log transform &1st difference to address **heteroscedasticity**;used **ACF/PACF** to select **ARIMA/SARIMA**
- Achieved **MAPE** value of **9%** with **ARIMA** model and performed **residual analysis** using the **Ljung-Box test**

Email Classification (NLP)
- Developed an email **classification** model to classify **spam & non-spam**, using a dataset of over **5.5k+** emails
- Trained models utilizing **Naïve Bayes & Logistic Regression** algorithm; used **TF-IDF** for **feature extraction**
- Enhanced performance with **LSTM**, resulting in an improved **Recall** value of **0.93** & **Precision** value of **0.97**

Image Captioning (Deep Learning)
- Built an **Image-Captioning** model using transfer learning with **ResNet-50** as encoder and **LSTM** as decoder
- Utilized **nucleus** and **temperature** sampling to build vocabulary using **NLTK** , achieving **ROUGE** score of **0.2**

Restaurant Recommendation System
- Developed **Content-Based** Restaurant recommendation system utilizing data extracted from **Yelp** data set
- Utilized **K-Dimensional Binary Tree** Structure and incorporated latitude and longitude for **localized** search
- Applied **TF-IDF & Cosine Similarity** for recommendations and validated results using **K- Means Clustering**

Intermittent Demand Forecasting
- Categorised the products into **Intermittent**, **Smooth**, **Lumpy** & **Erratic** based on ADI & CV^2 |Croston method
- Forecasted demand for intermittent series with **Syntetos** & **Boylan** method. ↓**MASE** to **1.67** from **4.14** (SES)

ADDITIONAL PROJECTS

Financial Insights with RAG and LLM (Gen AI)
- Built **Chatbot** using **Retrieval Augmented Generation** & **Langchain** to generate **Financial Report Analysis**
- Loaded docs and created chunks via **TextSplitter**; created embeddings using **OpenAI** & stored in **Pinecone**
- Used **MMR** for retrieving chunks; incorporated memory and created **retrieval chain** for efficient querying

Credit Risk Modeling (Classification)
- Developed a Credit Risk Model to **optimize bank credit lending strategy** using dataset of **50k+ customers**
- Executed **feature engineering** and **selection** using **Chi-Square** & **VIF** tests for categorical & numerical data
- Compared **DT, RF** and **XGBoost** classifiers; optimized XGBoost with **GridSearchCV**, achieving **78%** accuracy

Customer Segmentation (Clustering)
- Analyzed Indian **bank** customer dataset with over **1L+** observations to identify distinct **customer segments**
- Performed **feature engineering** and **RFM analysis** | explored **K-Means, Agglomerative, GMM** and **DBSCAN**
- Identified **7** segments using **elbow method**, did cluster profiling; **silhouette score: 0.54** and **DB index: 0.79**

Multi-class CNN Image Classification
- Built **CNN** model to help companies in classifying **101** classes of food products, from a dataset of **1L+** images
- Implemented **Data Transformation, Augmentation** and **Batch Normalization | Dropout** for regularization
- Used **transfer learning** with pre-trained **VGG-16** and **DenseNet121**, improving accuracy from **74%** to **80%**

AWARDS AND ACHIEVEMENTS

Competitive Exams
- Top **1%** out of **10 lakhs+** candidates to qualify **UPSC Civil Services Prelims** in 2018 , 2019 , 2020 and 2023
- Cleared UPSC **Indian Forest Services -** Prelims exam twice in 2018 & 2023 and UPSC **CAPF** Prelims in 2022
- Qualified for **Haryana Civil Services** Mains Exam (Top **1 %** of **1.5lakh+** candidates) in 2017,2021 and 2022

Case Competition
- Ranked **4th** (out of **900+ contestants**) in **IDB Analytics 3.0** competition, organized by **IIM Calcutta** in 2023

District Level Awards
- Secured **second** position in Sirsa **District Yoga Championship**, conducted by **Haryana government** in 2004
- **Bal Vigyanik Award** by **Haryana government** for study on impact of **BT-Cotton** in agricultural fields (2003)

POSITIONS OF RESPONSIBILITY & EXTRA CURRICULARS

Social Work at SOREM
- **Pro-bono** work for 6 months at **Society for Rehabilitation of Mentally Challenged**, Chandigarh in 2022-23
- Worked as **Assistant Teacher** for special **Children** & **Adults** with **physical & mental disabilities** at SOREM
- Implementation of **Smart Classes** tailored to unique requirements of students, for their holistic development

NSS, IIT Madras
- Member of IIT Madras's **National Service Scheme**, particularly in **Environmental Conscious Group** (2011)
- Undertook a **beach cleaning** initiative at **Kottapakkam** beach in **rural** district of Ongole, **Andhra Pradesh**
- Organized **Education Awareness Program** for underprivileged children for class - X preparation, in Ongole

Sports
- Secured **1st** position in **Badminton** Singles Tournament in Inter-Hostel competitions at **IIT Madras** (2012)

ELECTIVES : Deep Learning, Financial Risk Mgmt., Econometrics	INTERESTS : Current Affairs, Running