

KAPIL PUSHKAR



ACADEMIC QUALIFICATIONS

Year	Degree /Board	University /Institution	%/CGPA
2025*	Post Graduate Diploma in Business Analytics	IIM Calcutta, IIT Kharagpur, ISI Kolkata	-
2021	Dual Degree - B.Tech + M.Tech. Civil Engineering	Indian Institute of Technology, Varanasi	7.04/10
2016	CLASS XII	Macro Vision Academy	82.8 %
2014	CLASS X	Macro Vision Academy	9.4/10

KEY SKILLS/TOOLS Machine Learning, Python, Data Analytics, Data Visualization, SQL, Power BI, Predictive Modeling, Excel

INTERNSHIPS (2 Months)

Colliers International	Project Mgmt. Intern	Gurugram (May '19 - Jun '19)
Roles and Responsibility	<ul style="list-style-type: none"> ■ Supervised site, ensured adherence to schedules, and optimized project duration through risk management ■ Led stakeholder meetings, addressed progress gaps, and maintained project trackers for effective planning ■ Enabled Cross-functional Collaboration across multiple contractors which lead to reduce in project timeline 	

AWARDS AND ACHIEVEMENTS

Case Competition	<ul style="list-style-type: none"> ■ Ranked 4th out of 1218 teams in Zenith 2.0 2023, Business Case Competition organized by IIM Calcutta ■ Ranked 8th out of 760+ teams in Ad-War-Tize 5.0'23, Marketing Case Competition organized by IIM Raipur
Awards	■ Secured 3rd place in Bridge-it at Technex, the annual techno-management festival of IIT Varanasi 2018
Scholastic Achievements	<ul style="list-style-type: none"> ■ Ranked 2672 out of 1.8 L+ aspirants (Top 1.42 %) in UP State Engg. Exam Among Top 6.7% in JEE-Adv ■ Awarded cash prize by C.R Employee's Co-operative Credit Society for outstanding performance in HSC
Certificates	■ Certified in " Python Data Structures " and " Using Python to Access Web Data " from UMich via Coursera.

ACADEMIC PROJECTS

Customer Churn Prediction Model (Genetic Algo)	<ul style="list-style-type: none"> ■ Developed Neural Network Model for employee attrition using Genetic algorithm to tune hyperparameter ■ Implemented Exploratory Data Analysis, Label Encoding; achieved accuracy of 84.2% & F1 Score of 90.35% ■ Model outperformed Logistic Regression, Support vector Machine, Decision Tree, Random Forest & XGB ■ Improved accuracy by 2.64% & F1 Score by 2.67% over the best performing model on imbalanced dataset
Stock price Prediction (Regression)	<ul style="list-style-type: none"> ■ Regressed log of Amazon's stock returns (4 years of data) with CAPM, Fama-French 3 & 5 Factor models. ■ Tested models on 2 months of data, achieving a model with MAE of 0.0075 & Adjusted R² value of 0.519. ■ Verified Heteroscedasticity using Breusch-Pagan test and Autocorrelation using Durbin Watson test.
Sentiment Analysis (NLP)	<ul style="list-style-type: none"> ■ Scrapped over 30K mobile reviews from Flipkart; performed topic modelling & sentiment classification ■ Performed tokenization, lemmatization & POS tagging; used LDA; coherence identified 6 optimal topics ■ Trained LSTM using labelled dataset; achieved 90.4%accuracy; got topic & sentiment level for test reviews
Early Lung Cancer Detection (Classification)	<ul style="list-style-type: none"> ■ Designed an ensemble CNN model, enhancing diagnosis capabilities & robustness with data augmentation. ■ Trained & Integrated EfficientNet-B3, Inception-v3, & ResNet50 using transfer learning for enhancement. ■ Achieved 91.25% accuracy, 2.27% over the leading model, with 1.7% in recall & 2.55% in F1 Score.
Predictive Maintenance Using (XAI & ML)	<ul style="list-style-type: none"> ■ Preprocessed data with SMOTENC used one-hot encoding and standard scaling for robust model training ■ Reduced dimensions with PCA and applied Gradient seachCV to identify best model based on accuracy ■ Best model Gradient boosting achieved balanced accuracy of 79.27% and Macro F1 Score of 75.68%

ADDITIONAL PROJECTS

Sales Forecasting (Time Series)	<ul style="list-style-type: none"> ■ Forecasted drug sales based on 13 yrs historical data and checked stationarity of data using the ADF test ■ Analysed ACF & PACF plots for order of SARIMA model & Ljung-box for the residual autocorrelation analysis ■ Selected best model using gridsearch based on AIC&BIC, obtained MAPE of 8% using best SARIMA model
Customer Segmentation (Clustering)	<ul style="list-style-type: none"> ■ Segmented 1M+ ecommerce customers using Pyspark to enhance Targeted Marketing based on invoice ■ Performed RFM analysis, mitigated outliers and segmented customers into 5 groups using an elbow plot. ■ Used K-means, Agglomerative & DBSCAN for clustering, achieving best Silhouette score 0.53 with K-means
Portfolio Optimization	<ul style="list-style-type: none"> ■ Computed annual portfolio return and risk of 5 stocks (Disney, Microsoft, Amazon, Google, and Walmart) ■ Refined portfolio to maximize Sharpe Ratio(1.08) on the efficient frontier, optimized risk-adjusted return ■ Leveraged the pypfopt library to discretely allocate stocks for each asset given certain investment amount
Health Care Analytics	<ul style="list-style-type: none"> ■ Applied K-means and Single-linkage agglomerative clustering techniques to categorize various hospitals ■ Leveraged the elbow method to identify the optimal no. of clusters and achieved a silhouette score of 0.58

POSITIONS OF RESPONSIBILITY & EXTRA CURRICULARS

Basketball	■ Winner in Sectional-Wars : Won a knockout tournament at IIM Calcutta , competing against 10 teams .
Social Work	<ul style="list-style-type: none"> ■ Supported over 150 families with essential donations during the COVID-19 crisis, ensuring their well-being ■ Actively participated in tree plantation and cleanliness drives across our district, promoting sustainability

ELECTIVES : FRM, Health Care Analytics, Supply Chain Analytics **INTERESTS** : Writing, Photography, Gym