

Shubham Khode

Chicago, IL | (858) 305-7003 | skhode2@uic.edu | [linkedin.com/in/shubhamkhode](https://www.linkedin.com/in/shubhamkhode) | www.shubhamkhode.me

EDUCATION

University of Illinois Chicago

Master of Science in Business Analytics (GPA: 4.0/4.0)

August 2021 – December 2022

Veermata Jijabai Technological Institute (VJTI), Mumbai, India

Bachelor of Technology in Electronics & Telecommunications Engineering (GPA: 7.5/10)

August 2012 – April 2016

SKILLS

- **Technology:** Python (NumPy, Pandas, Scikit-Learn), R, SQL, C#, Hadoop, Spark, AWS, Tableau, Power BI, Excel, Git, Jira
- **Core Competencies:** Data Mining, Machine Learning, Predictive Modeling, Clustering, NLP, PCA, Descriptive & Inferential Statistics, Time Series Forecasting, A/B testing, ETL, Visualization, Agile Methodologies, DevOps

PROFESSIONAL EXPERIENCE

Wipro Limited

Business Analyst | Vienna, Austria

August 2019 – June 2021

- Conducted statistical analysis using SQL and R on customer's products & services data to discover underlying patterns for mitigating churn rate; made recommendations to bump up retention by 15%
- Designed ETL pipelines to fetch and manipulate data from disparate transaction database systems and consolidate it into a data warehouse for downstream BI reporting & analytics
- Applied A/B testing on different landing/product page variants of client's online shop; lifted conversion rate by 7% and marked increase of ~\$5M in revenue flow
- Identified borderline credit risk customers using ML algorithms; redesigned order workflows resulting in 2x faster lead time
- Built interactive Tableau dashboards for quick, actionable insights on core business KPIs; saved 8 hrs./week in manual reports
- Collaborated with key stakeholders and leveraged data-driven opportunities for aligning product roadmap to strategic goals

Software Engineer | Pune, India

August 2016 – July 2019

- Analyzed business requirements, documented high-level technical specifications, and developed E2E solutions while ensuring feasibility & integration with legacy systems
- Developed and enhanced 5+ business-critical web applications in .NET to provide a unified experience across all platforms
- Automated log management using Python and prepared dashboards to monitor application activity and performance; improved resolution time for production bugs by 35%, saving ~\$500K/year
- Implemented complex SQL scripts for applying pseudonymization; transformed 1M+ customer records with sensitive information using data masking procedures for EU GDPR compliance
- Explored historical sales & consumption data for a telco and utilized prescriptive analysis for optimizing price-plan portfolio based on levers such as data volume, speed, add-ons, and promos; observed ARPU uplift of ~3%

PROJECTS

Forecasting Retail Sales for Walmart at scale (*Python, Apache Spark, Facebook Prophet*)

- Analyzed influence of factors like markdowns, holidays, fuel price, temperature, etc. on weekly sales over 450K+ datapoints.
- Trained hundreds of time series models in parallel by leveraging Spark distributed data processing and generated precise forecasts for each store-department combination using Prophet library.

Loan Default Prediction and Investment Decisions for Lending Club (*R, EDA, Ensemble*)

- Performed EDA (univariate, multivariate, correlation) to understand loan characteristics and driving features behind defaults.
- Built Decision Tree, Random Forest, and GBM models to predict the probability of loan default and estimate expected annual returns. Evaluated performance using AUC, ROC, and lift curves and combined best models to maximize ROI.

Spotify – Top Songs & Artists Popularity and Recommendation (*Python, Tableau, Regression*)

- Investigated and visualized various acoustic features to answer riveting questions pertaining to song and artist popularity.
- Experimented with different regressors using grid search to predict popularity of songs. Designed a recommendation system based on neighborhood collaborative filtering with Cosine and Manhattan distances as similarity metrics.

Market Segmentation of Soap Company (*R, Clustering*)

- Executed K-means, K-medoids, and Hierarchical clustering algorithms on basis of purchase and behavior data. Profiled customers based on clusters identified from Elbow & Silhouette methods, proposed strategies targeting each segment.

Text Sentiment Analysis on Yelp Reviews (*R, NLP, Classification*)

- Applied TFIDF text vectorization on restaurant reviews and used dictionaries like AFFIN, Bing, and NRC to build classifiers for sentiment polarity. Achieved an overall F1 score of 0.92 with LASSO regularized Logistic Regression.