SOHAM SHINDE

Boston, MA 02120 • (857) 376-8612

shinde.so@northeastern.edu • linkedin.com/in/sohamsshinde • github.com/sohamthirty

EDUCATION

Northeastern University, Boston, MA — Aug. 2021 – Present
Khoury College of Computer Sciences — Expected Graduation: May 2023
Master of Science in Data Science — **GPA: 3.5/4.0**
Related Courses: Supervised Machine Learning, Data Management and Processing, Algorithms, Database Management Systems

University of Mumbai, Mumbai, India — Aug. 2017 – June 2021
Bachelor of Engineering in Computer Engineering — **CGPA: 9.4/10.0**
Related Courses: Big Data & Analytics, Natural Language Processing, Artificial Intelligence & Soft computing

TECHNICAL KNOWLEDGE

Programming Languages:	Python, R, SQL, Scala, Java, C, C++, PHP, HTML, CSS
Libraries:	Pandas, NumPy, Sci-kit, Keras, TensorFlow, NLTK, Matplotlib, Plotly, ggplot2,
Machine Learning Algorithms:	Regression, SVM, K-Means, Random Forest, Neural Networks, Ensemble Methods, Time-series forecasting, PCA, Boosting, Bagging, Sentiment Analysis
Tools:	Tableau, Power BI, Git, Hadoop, Hive, Pig, MapReduce, AWS, MS Office
Achievements/ Certifications:	Scrum Fundamentals Certified (SFC), Winner - ACM Mumbai Hackathon 2020
Publications:	ML-Based Smart Shopping System with Recipe Recommendation (ICCICT)

PROFESSIONAL EXPERIENCE

TwinTring LLP, Mumbai, India — Mar. 2020 – May 2020
Android Developer (Java, UI/UX, Android Studio, XML, APIs, Firebase)
- Developed GPS navigation app for tracking ride time, distance, calories burned, avg/max speed, elevation gain for riders, and analyzed monthly customer usage data to recommend personalized routes and identify gaps in cycle supply services.
- Programmed group ride feature to connect 10+ riders within 5km radius to share live location and connect via chat that increased the Maximum Distance Operation connectivity by 40%
- Improved the Crash-free Rate (CFR) to 98% by debugging the app for a focus group of 50+ users consisting of long distance, and off-terrain riders that reduced the latency from 3 to 1 sec

Teknack Gaming Studio, Mumbai, India — Jan. 2020 – Mar. 2020
Game Developer (Unity, Blender, C++, C#, SQLite)
- Collaborated with a cross-functional team of developers, designers, and testers to create an adventure-themed game by implementing 2D physics, sound engines, kinematics, and collision detection in Unity, and Adobe After Effects
- Achieved average rating of 4.9/5 on Play Store with 1000+ downloads within 30 days; awarded the Most Popular Game among 16 teams with a surveyed Customer Satisfaction Score (CSAT) of 82%
- Conducted usability tests on 20+ users that led to improvements in frame rendering, pixel rendering and Battery consumption; increased net promotor score (NPS) by 15%

PROJECTS

Electricity Price Forecasting (Python, LSTM, ARIMA, Time series, Neural Nets) — Oct. 2021 – Dec. 2021
- Forecasted the hourly electricity price for 12 months using time series data by identifying seasonal changes in 23 features of generation, consumption, and weather data
- Discovered trends and potential relationships by EDA, implemented feature engineering and hyper-parameter tuning for LSTM prediction model that reduced RMSE score by 32% compared to the existing result model

Real Estate Prediction (Python, Decision Tree, SVR, Light-GBM, AWS, Flask) — Dec 2021 – Jan 2022
- Predicted the selling price of properties with R2-Score of 82.43 using regression models (Linear, Decision Tree, Support Vector, Light-GBM) for comparative market analysis
- Delivered inference requests from user inputs by deploying the pickled model on AWS EC2 instance

ML-Based Shopping System with Recipe Recommendation (Python, NLP, TF-IDF, CBFA, LDA) — Oct.2020 – Mar.2021
- Created a recommender system that provided top 10 recipes based on ingredients in cart using Text Rank, Content-Based Filtering and Natural Language Processing to generate tags, and suggest products by collaborative filtering
- Conducted sentiment analysis by analyzing user feedback of 1500+ shopping transactions collected via mobile app

Bike-Sharing in Boston (R, R-Studio, ggplot, ggmap, dplyr, tidyverse) — Sept. 2021 – Oct. 2021
- Performed Data Wrangling on 4M+ records, filtering outliers, pivoting tables, handling messy data, and fixing structural errors
- Visualized the Trip Duration, Popular Stations, Age Wise-Trip Distance, Customer Type using ggplot and ggmap to interpret the impact of pandemic on bike-rides, shrinking subscription, and increased trip duration