

# VINAY SISODIYA

## Contact

Email ID: vsisodiy@andrew.cmu.edu

LinkedIn: [www.linkedin.com/in/vsisodiy](https://www.linkedin.com/in/vsisodiy)

Ph: 646-932-6687 | Add: Squirrel Hill, Pittsburgh, PA



## Education

**Carnegie Mellon University** | Master of Information Systems Management – Business Intelligence and Data Analytics Pittsburgh [Aug'19 – Dec'20]

- Courses: Intro to Deep Learning, OOP in Java, Machine Learning with Large Datasets, Database Management, Distributed Systems
- Graduate Teaching Assistant: 95828 - Machine Learning for Problem Solving, 18540 - Rapid Prototyping of Computer Systems

**Indian Institute of Technology - Bombay (IIT-B)** | Mechanical Engineering

Mumbai, India [Jul'12 – Apr'16]

- Courses: Data Analysis & Interpretation, Probabilistic Models, Decision Analysis and Game Theory, Linear Algebra, Economics

## Work Experience

**American Express** | Assistant Manager

Gurgaon, India, [3yrs | Jul '16 – Jun '19]

**CM Attribute Tool** | Business Insights | *Combined advanced analytics & behavioural economics to rank customers on lifestyle attributes*

- Scored customers on Price Sensitivity, Online/ Mobile Affinity, & staycation appetite using Gradient Boosting technique for regression
- Applied Tf-Idf statistic for industry weightage; Attained 70+ Gini coefficient across all 3 models; Customised offers for 10+ global merchants

**Commercial Matching Framework** | Customer 360 | *360° view of commercial entities utilised by other verticals like Fraud, New Accounts*

- Devised a framework using text-matching on firmographic attributes & Random Forest for predicting final match score using HIVE
- Created possible pairs using Map-Reduce & final linkages using Connected Components; Devised arbitration-rules to assign IDs to new accounts

**Car Buyer Model** | Business Insights | *XG-Boost Model to predict the potential Car buyers in US market*

- Achieved ROC-AUC of 83% using XG-Boost (explored Logistic Regression & SVM) using SQL; Model used for 3 global car manufacturers

**BAU Work:** Competitive Landscaping, Prospect Identification (KNN), Loyalty Assessment, B2B Merchant Recommender (Collaborative Filtering)

## Academic & Independent Projects

**Product Identification, Walmart** | Graduate Research Assistant | Biometrics Centre, Cylab, Prof. Marios

CMU - [May '20 – Aug '20]

- Devised complete pipeline from data cleaning using cosine similarity - designing multiple anchors – creating bounding boxes for label & product detection using RetinaNet on aisle panos to identify out of stock, wrongly placed, & wrongly labelled products; Achieved 95% precision & recall
- Utilised Nvidia Apex mixed precision for switching between FP32 & FP16 computations to achieve 2.2X training speed with similar precision & recall

**Listen, Attend & Spell: Attention Language Model** | Prof. Bhiksha Raj

CMU - [Mar '20 – Apr '20]

- Developed a Machine Translation/LAS model using cascaded LSTM networks for speech to language transcription on Wall Street Journal Dataset
- Deployed Beam Search Decoder (from scratch) & CTC Decoder to predict the phonemes in utterances achieving Levenshtein score of 8.27
- Modelled an attention-based LSTM transducer that generates a distribution over the next character conditioned on all previous characters along with Pyramidal Bi-LSTM speech encoder to reduce computational complexity; Obtained Levenshtein score of 7.92

**Paper Question Answering** | Prof. Bhiksha Raj

CMU - [Mar '20 – Apr '20]

- Implemented a novel approach of BiDAF context-query attention & modelling layers on top of SciBert embeddings for QA on academic paper abstracts
- Answered one word, number, short & long answer questions; Achieved 6% improvement in Top 1% error over Bert and DistilBert Baselines

**Face Classification & Identification** | Prof. Bhiksha Raj

CMU - [Jan '20 – Feb '20]

- Implemented ResNet-50 architecture from scratch using PyTorch to classify images of faces (2300 distinct) with an accuracy of 79%
- Utilised penultimate layer features to calculate cosine similarity between 2 images to check if they belong to same person with 93% accuracy

**Machine Learning Algorithms (From Scratch Implementation)** | Prof. Matt Gormley

CMU - [Aug '19 – Dec '19]

- **Sentiment Analyzer:** Classified a movie review with 94% accuracy using binary logistic regression over dense bag-of-word representation
- **HMM:** Learned prior, transition & emission probabilities using MLE; Utilized Viterbi algorithm to predict hidden states with 93% test accuracy
- **Decision Tree:** Trained classifier using ID3 splitting & Cross Entropy loss; Predicted politician's party based on Bill-Vote history with 86% accuracy
- **Mountain Car:** Obtained optimal policy & action values using Q-learning; Employed epsilon-greedy method & tiling representation of state space

## Skills & Expertise

**Languages & Libraries:** Python, Java, SQL, SAS, Hive, PyTorch, AWS, Spark, Map Reduce, Scikit, Numpy, Pandas, Spacy, Docker, Heroku

**Data Science:** CNN, Object Detection, NLP, LSTM, Decision Trees, Logistic Regression, Graphical Models, SVM, PCA, t-SNE, KNN, K-Means

## Miscellaneous

- AmEx Accomplishments – Won Analyst of the Quarter and Star Performer Award; Organized workshops and leader talks as part of L&D team
- Academic Accomplishments – CFA Level 1 certification (Jun'18); All India Rank of 99.66 percentile in JEE 2012 (out of 0.5M students)
- Leadership & Cricket – Served as Institute Cricket Secretary, IITB (Apr'14 – Mar'15); Gold Medal, Cricket Championship (Inter IIT 2014)
- Interests & Hobbies – Intrigued by sports analytics, Mentoring, Cooking, playing & watching Cricket, Adventure sports