Vaibhav Beohar

vbeohar@berkeley.edu ● 437-343-7743 ● Toronto ON ● linkedin.com/in/beohar/ ● beohar.com

Hands-on technology professional with experience in software engineering (front-end and back-end), data science, finance & accounting. Experienced practitioner with proven record in design, development and delivery of complex software systems and machine learning models.

SKILLS

- **Languages:** Python, Java, C++, SQL
- **Front end scripting languages:** HTML5, JavaScript, CSS, jQuery, AJAX, D3.js, React
- **Libraries:** TensorFlow, Keras, PyTorch, sklearn, NumPy, Seaborn, pandas, Spacy, Transformers
- **Applications and tools:** MongoDB, MySQL, Oracle, Snowflake, PostgreSQL, SQLite, Nginx, Gunicorn, Anaplan, Docker, Visual Basic, Google Cloud (GCP), AWS, GitHub, Heroku, MS Office, Atlassian Jira, Trello, Jupyter Notebook, Visual Studio, PyCharm, Sagemaker, Lambda
- **Competencies:** Full stack development, Design Patterns, Object Oriented Programming, Agile with Scrum, Exploratory Data Analysis, Machine Learning, Natural Language Processing (NLP), Deep Learning, Data Engineering, Quantitative methods for statistical analysis, experimental design and causal inference, Financial Modeling, Risk/Planning & Analysis, Corporate Finance, Stakeholder management, Business Analysis, Visualization, Bayesian modeling, A/B testing

EXPERIENCE

MCKINSEY & COMPANY *Waltham MA/Toronto, ON*
Data Scientist & Finance Specialist *March 2017– ongoing*

- Created unsupervised information extraction NLP model using Python and Google's Universal Sentence Encoder on TensorFlow to implement an internal *search engine* for service line segmentation, advertising, client conversion and classification using media articles via Python, Snowflake and exploratory web technologies like Flask, React, jQuery and D3.js
- Developed and refactored random forest and logistic regression based supervised learning models on R Studio and Python to generate predictive probabilities on client churn, win, winback and service line segmentation, with 90% AUC
- Partnered with data engineers to automate REST API based extraction, transformation and load (ETL) jobs and wrote scripts to create Python, S3, PostGreSQL and Snowflake based multi-threaded data engineering pipelines reducing load times by 5x
- Collaborated with cross-functional teams on implementing a unified snowflake-based *feature store* for machine learning use-cases by conducting architecture reviews and providing recommendations on best practices; thereby eliminating Python based ETL code redundancy, bringing economies-of-scale and reducing model run time by 25%
- Automated firm's annual budgeting and forecasting process , spread across 50+ countries and 500+ users, into Anaplan (a cloud based EPM platform); by streamlining headcount, workforce utilization, opex, balance sheet, income statement and other miscellaneous KPI planning – thereby improving process efficiencies and user engagement by 250%
- Supervised, coordinated and lead discussions on various ad-hoc analytic and data science discussions and prospective engagements using project management skills, Agile (with scrum) methodologies and tools such as Jira

GLOBAL ATLANTIC FINANCIAL GROUP *Boston, MA*
Associate – Risk, Planning & Analysis *May 2012– March 2017*

- Prepared finance materials for CFO, monthly operating committee meetings, quarterly board and risk meetings and other ad-hoc projects, with timely analysis of various business lines for corporate finance decision making
- Automated legacy Excel models using VBA Macro and SQL for planning, budgeting, and forecasting of short-term and long-term net income of 8 sub- and 1 consolidated entity ($52bn assets and >300mm operating income); leading to 95% reduction in execution/delivery time compared to prior forecasting models
- Overviewed Actual-to-Expected (A-to-E) variance analysis by cross-functional collaboration with controllers, actuarial, tax and expense teams and by mapping model-to-ledger income statement and balance sheet accounts
- Identified market, regulatory, accounting, and operational items impacting excess capital availability, including:
    - What-if scenario analysis with low/high/baseline test cases and KPIs reported on capital, OpEx and CapEx
    - Risk charges owing to business, interest rate, asset risks, reinsurance, hedges and derivatives
    - $2bn+ Capital attribution analysis by new business, new sales, net investment income (NII) etc.

GOLDMAN SACHS *New York, NY*
Senior Consultant Programmer *April 2007– Aug 2010*

- Developed client-side portfolio management portal (https://marquee.gs.com/) and various trading tools for Goldman Sachs Asset Management (GSAM) using Java, Unix, Perl, shell scripts, Sybase SQL and various open-source technologies
- Designed and developed a real-time trading solution, as well as an advanced risk model, with timely rollout during the critical Troubled Asset Relief Program (TARP) to assist the fixed-income trading desk with mortgage-backed security trading, portfolio tracking and performance monitoring tasks using IT tools
- Decommissioned and replaced legacy systems with consolidated web-based system for fixed income insurance desk
- Interacted with external vendors (such as Interactive Data BondEdge) for integrating 3rd party applications for GSAM portfolio management and reporting needs
- Developed, tested, and supported core modules for rollout of GSAM's fixed income insurance advisory system

SIGMA SOFTWARE SOLUTIONS *Pune, India / Toronto, ON*
Senior Software Engineer *July 2004– Dec 2007*

- Lead 4-member offshore team on various order, revenue & customer management enterprise level technology initiatives using advanced Java, J2EE, Unix, MySQL, open-source tools (CVS, Apache Struts, JUnit, Hibernate ORM)
- Developed, researched, and documented technical project artifacts and software design documents
- Lead a "first in company" novel proof of concept of an end-to-end enterprise web-based system using Hibernate object relational mapping framework for an Africa based telecom client. Customer was highly appreciative of the effort and was engaged in long-term contract
- Underwent a 4-week companywide intensive bootcamp on advanced Java. Stood 1st out of 35 trainees

EDUCATION

University of California Berkeley; School of Information *Berkeley, CA*
Master of Information & Data Science (**MIDS**) – GPA 3.97 *August 2022*

- Coursework – data engineering, statistics, applied machine learning, causal inference & experiments, deep learning on edge devices, natural learning processing with deep learning, data visualization, capstone
- Projects:
    - Performed Airbnb price prediction using XGBoost, AutoML, neural network & logistic regression
    - Trained deep learning models on IBM cloud and performed inference on Nvidia Jetson TX2 edge devices
    - Created full-stack website using D3, Flask and Python to display graph visualizations for Star Wars movie franchise
    - Explored Aspect Based Sentiment Analysis (ABSA) techniques and implemented fine-tuning of RoBERTa, SpanBERT and DistillBERT models to achieve material improvements in aspect extraction and classification tasks
    - Capstone project – Created an end-to-end content recommendation analytics product called "MakeSense" to allow emerging streamers on Twitch expand their presence on the platform using BERT, HuggingFace, D3.js and Flask

University of Massachusetts Amherst; Isenberg School of Management *Amherst, MA*
Master of Business Administration (**MBA**) Finance & Accounting – GPA 3.82 *May 2012*

- Internship (June - Aug 2011): Goldman Sachs – Investment Accounting/Controllers (Jersey City, NJ)

Rajiv Gandhi Technological University *Bhopal, India*
Bachelor of Engineering in Computer Science (Honors) – GPA 3.75 *June 2004*

CERTIFICATIONS

- Chartered Financial Analyst (level I) and "Official Anaplan Developer" certification