Vaibhav Beohar

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

Skilled professional with experience in software engineering (front-end and back-end), data science, finance & accounting. Currently working as a data scientist at McKinsey. Obtained an MBA in 2012 and master's in data science from UC Berkeley

SKILLS

Languages: Python, Java, C++, SQL
Front end scripting languages: HTML5, JavaScript, JSON, CSS, jQuery, AJAX, D3.js, React.js
Libraries: TensorFlow, Keras, PyTorch, sklearn, NumPy, seaborn, pandas, Spacy, Transformers, Java Enterprise Edition (J2EE)
Applications and tools: MongoDB, MySQL, Oracle, Snowflake, PostgreSQL, SQLite, Nginx, Gunicorn, Anaplan, Docker containers, Bootstrap, Visual Basic, Google Cloud, AWS, DataBricks, GitHub, Heroku, MS Office, Atlassian Jira, Trello
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, Financial Modeling, Risk/Planning & Analysis, Corporate Finance, Stakeholder management, Business Analysis

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 JavaScript, Bootstrap, Flask and Python to display "social network" graph visualizations for the Star Wars movie franchise using D3.js
  - 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.
    Technologies: Python, Flask, PostgreSQL/SQLite, jQuery/Ajax, D3.js, Nginx and Gunicorn, BERT model double fine-tuned on Masked Language Modeling (MLM) and Sentiment Classification tasks

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*

EXPERIENCE

MCKINSEY & COMPANY *Toronto, ON*
Data Science Specialist *March 2017– ongoing*

- Created an unsupervised information extraction NLP model using Python and Google's Universal Sentence Encoder on TensorFlow and Nvidia K100 GPU to implement an internal *search engine* for McKinsey's client documents and match them with prospective service line offerings using Snowflake, Python, Flask, jQuery and D3.js
- Performed monthly run of random forest and logistic regression based supervised learning models on R Studio and Python to evaluate McKinsey's 2000+ active/lost clients and create distribution of *most likely* service line offerings with 90% AUC
- Built asynchronous Flask server based natural language processing (NLP) model for sematic information extraction and use-case

ranking using MongoDB, Word2Vec, EventRegistry API and Python. Used Python multiprocessing code and a multi-compute environment to reduce latency by 80%

- Worked on machine learning models to predict headcount attrition, workforce monthly utilization, regional opex targets, cash reserves and various other key metrics; and integrating the results into the annual budgeting/planning software, Anaplan (a cloud based EPM platform) for firmwide distribution and user engagement
- Supervised, coordinated and lead discussions on various ad hoc data science discussions and prospective engagements using project management skills, Agile (with scrum) methodologies and tools such as Jira
- Key technologies and skills – Python, Jupyter Notebook, R Studio, regression and supervised learning, SQL, NLP, MongoDB, PostgreSQL, word2vec, BERT, Spacy, jQuery and D3.js, Ubuntu, shell scripting

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 (http://360.gs.com) and various trading tools for Goldman Sachs Asset Management (GSAM) division using advanced 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

CERTIFICATIONS

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