Anurag Bejju

Data Scientist



+1 (778) 316-5973

Vancouver, Canada



WORK EXPERIENCE / CO-OP

Data Scientist Intern

May 19 - Present

Statistics Canada, Ottawa

• Built a state-of-the-art solution that automated the entire process of extracting key data points from a really large range of companies financial documents (close to 2TB) with significantly varying structure and formats using the power of Natural Language Processing, Page Clustering, and Graphical Text mining techniques. - (Link)

Application Support Engineer

Jun 16 – Apr 18

ROAMWORKS LLC, Dubai

- Developed a Server Management Portal that monitored various performance metrics providing stakeholders a one-point solution to monitor nineteen production server in real time.
- Worked on multiple development projects, revamped existing legacy ROAM mobile application and worked on big data tools like Pentaho to design ETL Jobs that can effectively transform and analyze large volumes of data.

Professional Services Intern

Mar 16 – Jun 16

Thomson Reuters, MENA

Developed an Online Project Management Portal used by 55 project managers and 120 consultants at Thomson Reuters MENA Region. It contained live BI Dashboards highlighting vital project information.

Microsoft Student Partner

Dec 15 - Dec 16

Microsoft Gulf, Dubai

 Conducted workshops, hackathons, tech events as well as worked with Microsoft professionals during technical and training events.

2

ADDITIONAL WORK EXPERIENCE

Teaching Assistant

Sept 19 - Present

CMPT 165: Introduction to the the WWW - Simon Fraser University

 Provided assistance with assignments and lectures to 126 undergraduate first year students.

Teaching Assistant

Jan 19 - Apr 19

CMPT 127 - Computing Laboratory - Simon Fraser University

 Provided assistance with C programming lab assignments to 120 undergraduate first year students.

Teaching Assistant

Feb 15 – Jun 15

Chemistry Lab - BITS Pilani Dubai Campus

 Conducted lab experiments as well as provided support to students to perform them successfully.



EDUCATION

Masters in Computer Science, Big Data

2018 - Present

Simon Fraser University

Location: Vancouver, Canada CGPA: 3.9 (Out of 4.33)

B.E. (Hons) in Computer Science

2012 - 2016

Birla Institute of Technology and Sciences, Pilani

Location: Dubai, United Arab Emirates

CGPA: 8.57 (Out of 10)



Data Science - Languages and Frameworks: Hadoop, PySpark, MapReduce, PyPy, Octave, NumPy, Pandas, Cassandra, SKLearn, PyTorch, Web Scrapping, NLP (Word2Vec,Genesis, Vader)

Big Data Systems*: Openshift, Docker, Kubernetes, AWS, Nginx, Gunicorn, Microsoft Azure, Supervisor, Heroku [*Basics]

Programming Languages : Python, C, SQL, Java, Smalltalk, JQuery, Flask, HTML5, PHP, CSS3, ASP.NET, IIS, HTML.

Visualization Tools : Matplotlib, Chart.js, Apex.js, Sparkline.js, D3.js, Tableau

OS: Linux, Windows, Mac OS.

Certifications: Machine Learning – Coursera – Andrew Ng, Deep Learning Specialization* (In progress)



ACADEMIC PROJECTS

TradeSpade: Price Signal Forecast for Financial Assets

Jan 19 - Apr 19

- Developed a solution that provides assistance to day traders with intraday trading by predicting Buy and Sell signals in order to maximize profits and make optimized decisions. Also measured the influence of social media and everyday news on market fluctuations.
- Collected 9.71 million records from 6 sources, performed EDA and used XGBoost Algorithm to predict price signals. Finally developed a dashboard using Flask and hosted it on cloud.
- Link: Web Application

Design Analysis of WePay's Highly Available MySQL Architecture - (Link)

Jan 19 - Apr 19

Performed a detailed analysis of WePay's MySQL distributed database architecture and assessed their components, design tradeoff's, failover mechanism, architecture, scalability and availability limitations, and other metrics.

The Battle for Bookings: A Comparative Assessment of Hotels and Airbnb's on a Global Scale - (<u>Link</u>)

Sep 18 - Dec 18

- Compare the costs of hotel and Airbnb rates in 88-cities around the world by leveling the measuring parameters to see how the deals stack up.
- Used python libraries like GeoPy, GeoPandas, NumPy to transform 2.4 million*100 records to match Airbnb to the nearest hotel and perform data analysis with visualization.

GIF-based Communication Using Sentiment Analysis-(<u>Link</u>) Sep 18 - Dec 18

• Given a text message inputted by the user, our project aims to identify an appropriate GIF that understands the context and sentiment of it. We processed 125,000 GIFS to get description and emotion and mapped it with the input text's emotion and value using semantic textual similarity.



SELF-DIRECTED RESEARCH AND PUBLICATIONS

Real Time Banking Solutions: Era of Cashless Payment Dec 16 **Systems**

- Provided an overview of real-time banking systems and proposed an implementable model which can eliminate the negative perception related to cashless payment systems.
- It also elaborates the growth potential of this technology and explores the innovative applications of real-time systems in banking institutions.

Recipe-Based Recommendation Systems: Using Content-Based Filtering Algorithm

Proposed a goal-oriented recipe based recommendation system using machine learning methods and information retrieval techniques.

Sales Analysis of E-Commerce Websites Using Data Mining **Technique Jan 16**

- Presented a technique to evaluate, propose and improve traditional pricing strategies by using web mining techniques to collect information from e-commerce websites and apply data mining methods to induce and extract useful information out of
- This paper was accepted by International Journal of Computer Applications, USA and was published in their January 2016 edition.
- It was also presented at The Third United Arab Emirates Undergraduate Student Research Competition, (UGSRC) held at Abu Dhabi University and was widely appreciated the judges. (Link)

Application of Information Technology in Accounting Management Systems Nov 15

- Established the effect of different information technology software applications for varied difficulty level to collect, store, retrieve and maintain accounting information.
- This paper was accepted for the Third International Conference on Business and Management in Connected Era held at Skyline University College,

Analysis of E-Commerce Strategies Based on Online Consumer Shopping Behavior

- Evaluated various strategies used by online giants like Alibaba, Amazon, Ebay, Rakuten as well as comprehend the access price sensitivity of online consumer.
- It was accepted by International Journal of Arts and Sciences (IJAS), Harvard University, USA.

Cloud Computing for Banking and Investment Services

Nov 14

- Proposed a model that can effectively integrate cloudcomputing techniques into the current frayed banking structure and eliminate negative perception regarding cloud computing in the banking industry.
- It was accepted and published by World Congress on Management of Accounting, Finance, Marketing, Entrepreneurship, Economics and Business (MAFMEB - 2014) conference held at Jawaharlal Nehru University, New Delhi and was published in the November issue (Link).
- It was also presented in Technofest'14 and received the Best Technical Paper Award.

HACKATHONS AND TECHNICAL COMPETITIONS

VanHacks - Vancouver

- •Were among the 8 finalists out of 26 teams at #VanHacks2019 hackathon held at BCIT - Vancouver.
- Our app Trash'Bae was a web-based educational game that will help the next generation across BC recycle better. (Link)

Hacking Health - Ottawa

2019

- Secured the "IBM Ottawa Design Studio" prize as well as the "Most Patient focused" award at the 3-day #HackingHealth Dual city hackathon held at Shopify Ottawa
- Developed an Alexa app that helped elderly people with dementia follow a routine using sensor technology. (Link)

Enginuity'16 2016

- Secured second prize in a coding competition called Code Blitz where 16 programs in different languages had to be decoded.
- Secured second prize in a coding competition called Bits and Bytes where encrypted questions had to be solved using python.

Technofest'15 2015

 Won second prize in Code.Compile.Run() conducted by CodeChef, where a problem needs to solved within a time frame.

Technofest'14 2014

- Secured second prize in Management paper presentation on Cloud computing in banking and investment services.
- Secured third prize in Mechanical Event called Angry Bird were catapult must be developed to demolish a pyramid of cups.

Enginuity'13

Secured first prize in Mechanical Event called Home Alone were a working trap must be designed in order to catch thieves.

Technofest'13 2013

Secured second prize in Mechanical Event called Poseidon were a working model of a boat must be designed out of cardboard.

EXTRA CURRICULAR ACHIEVEMENTS

Red Bull Carve the Can Competition

2014

2013

• Received third prize and won 500 AED cash prize for designing the exterior of my red bull can.

Artex'14 2014

- Received first prize in Paper Mosaic where a portrait must be made using recyclable material.
- Received first prize in Rainbow Veins, a two people relay race consisting of 7 tasks to be completed.
- Received second prize in Classroom Art, an art form where your drawing needs to be abstract in form.

2013 Artex'13

- Received first prize in Classroom Art, an art form where your drawing needs to be abstract in nature.
- Received second prize in Envisage were we had to design an innovative cover page for chemical department magazine, CheSa.
- Received third prize in Pen Art, an art form were you are only allowed to use pen to showcase creativity.
- Received third prize in Clay Do Scope were we had to make a cartoon character using clay.

Red Bull Race the Can Competition

2012

 Received second prize in University level and was qualified to represent my college in UAE level held in Abu Dhabi's Corniche Beach.