

# MOHAMMAD ARBAZ KHAN

Arbaz1khan786@gmail.com | +1 289 885 3178 | <https://www.linkedin.com/in/arbazanalytics/>  
<https://github.com/ArbazAnalytics>

## PROFILE

Data Scientist Intern with a post graduation certification in Big Data Analytics program and working experience with data (extraction, entry, validation, formatting, importing), data classification, Large Language Model, Natural Language Processing

## EDUCATION

**Post Graduation in Big Data Analytics, Lambton College** May 2022 – Dec 2023 | Toronto, Canada  
**Bachelors in Electronics and Communication Engineering, IPEC** Jul 2016 – Oct 2020 | Ghaziabad, India

## SKILLS

**Programming Tools and platforms** (SQL, Python(NumPy, scikit-learn, Pandas, PyTorch, Transformers, Matplotlib), Pine Script, MySQL, Trading View, Microsoft Power BI, Tableau, JSON, Microsoft Office (Word, Excel, PowerPoint, Outlook, OneNote), Github, ChatGPT, Cohere, Together AI, Mistral AI, AWS RDS, Visual Studio Code)

**Soft Skills** (Problem solving, analytical Skills, organizational, time management, analytical thinking, attention to detail, teamwork, critical thinking, communication, sales, interpersonal)

## PROJECTS

**JobHelp** Jan 2024 – present  
- Ideated and collaborated with team, to design database for storing information of 99% of resume  
- Developed a platform for storing students' resume, established database on Amazon Relational Database Service and connected it utilizing MySQL

**The GenAI Competition, Hackathon** Dec 2023 – Jan 2024  
Built idea validating tool called 'Idea Check' for saving 80% time of founders on selecting ideas sent by entrepreneur

## PROFESSIONAL EXPERIENCE

**Data Scientist Intern, StableHacks** Sep 2023 – present | Mississauga, Canada  
- Eliminated manual data extraction process through development of a Python automation code  
- Automated by leveraging libraries such as Beautiful Soup, and OpenAI API to scrape data on funding programs from over 1600 finance websites, connected through MySQL  
- Planned and designed database by understanding finance websites data on MySQL  
- Build OpenAI prompt for 16 columns for obtaining exact data and fine-tuned it for improved accuracy, cleaned using SQL and Power BI  
- Categorized over 1200 distinct rows of a column named 'program target' into 10 main categories by utilizing advanced SQL operators ('CASE', 'LIKE', 'CONCAT')  
- Engineered a comprehensive bag of words table to enhance data retrieval by 90% within chatbot  
- Reduced cost of data extraction to 50% by leveraging Together AI API and the Mistral model  
- Trained Cohere Classifier, for positive or negative feedback on a review, fine-tuned it to 99% accuracy, by providing equal number of labelled examples, used Cohere API

**Stock Analyst Intern, StableHacks**  
- Analyzed candle stick pattern on trading view website by creating Tweezer, Hammer, Hanging Man indicators through Pine script programming language for up to 100% profit  
- Enhanced Turtle Soup, Tweezer, Breakout, Moving Average Crossover trading strategies up to 100%

## AWARDS

**BlockHack** 2023  
Won BlockHack Hackathon Based on Blockchain used Solidity programming, MetaMask, HTML, Python for creating a solution for storing medical history of patients on a centralized platform

**IIT JEE Exam** 2016  
Cleared IIT JEE main exam, a well recognized in India to get admission in Engineering College

## ORGANIZATIONS

**Regeneration Outreach Community, Cloth organizer** Apr 2022 – Sep 2022 | Brampton, Canada  
Categorizing a diverse range of 20+ cloth types at thrift store, optimizing inventory management for seamless operations