deepakchawla35@gmail.com   deepakchawla.me

+91-969-444-2980

Deepak Chawla

OBJECTIVE: I aim to join a dynamic organization and utilize my skills in evaluating and adopting new technologies, managing technology budgets, and collaborating with cross-functional teams to drive innovation and achieve business objectives. I have 6+ years of experience in building data-intensive applications and overcoming complex architectural and scalability challenges across diverse industries. I am proficient in Python, predictive modelling, computational linguistics, data processing, text mining, and data retrieval algorithms. My capabilities include creating, developing, testing, and deploying highly adaptive services that effectively translate business and functional requirements into substantial deliverables.

SKILLS

- **Languages**: C, C++, Shell Scripting, Python, SQL and PySpark.
- **Databases**: PostgreSQL, MySQL & MongoDB, Cosmos, HBase, HDFS and Cassandra
- **Data Analysis and Visualization Libraries** - NumPy, Pandas, Statsmodels, Tableau, PowerBI, Dask, Matplotlib, Plotly, ETL, D3
- **Machine and Deep Learning Libraries:** Sklearn, Tensorflow, Caffe & Keras.
- **Computer Vision and NLTK Libraries**: NLTK, Hadoop, Spacy, HuggingFace, Gensim, Opencv & Pillow
- **Machine Learning Algorithms**: Supervised, Unsupervised Algorithms (Regression, Random Forest, Decision Tree, SVM, SVC, KNN, K-means, Naive Bayes, Xgboost), Dimensionality Reduction (PCA), NLP, Word2vec, USE, Bert, LSTM Generation, Text Classification, Statistical modelling, sales forecasting and model optimization.
- **DevOps:** Docker, CI/CD using Jenkins & Kubernetes, MLFlow
- **Deployment** - Flask/Fastapi, Tensorflow serving and Lambda functions, Azure Functions, Azure Synapse, Azure Kubernetes Services
- **Cloud Platform:** Azure Cloud, AWS Cloud, Google Cloud, Data Engineering using AWS like Glue, Redshift, Databricks, Sagemaker Studio, Jupyterlab, VSCode, Pycharm, Apache airflow, Spark-Streaming, Storm, Event Hub, Kafka
- **Other**: Curl, Git, DVC, Beautifly Soap,

PROFESSIONAL EXPERIENCE

CoffeeBeans, Bangalore   **(6 months - Sep 2022 - Feb 2023)**
Lead Data Scientist   (Gen AI, ML, DS, Team Management, MLOps)

- As an accomplished Lead Data Scientist, I am actively spearheading cutting-edge initiatives, focusing on the development of powerful LLM (Large Language Model) models, GEN AI, and an in-house GPT (Generative Pre-trained Transformer) system. My proficiency in advanced analytics and machine learning is instrumental in implementing MLOps practices to enhance model deployment, monitoring, and optimization.
- In addition to my AI research contributions, I lead strategic efforts in building robust data warehousing solutions and architecting innovative system designs. My technical acumen enables me to streamline data management processes, unlocking valuable insights from vast datasets and translating them into actionable strategies.
- Taking a hands-on approach, I foster a culture of innovation and collaboration within my teams, ensuring they excel in delivering impactful data solutions. My passion lies in creating dynamic frameworks that facilitate continuous improvement and drive exceptional results.
- Let's connect to explore how my comprehensive expertise in data science, machine learning, and technical prowess can elevate your organization to new heights. Together, we can harness the full potential of data-driven solutions to achieve unrivaled success.
- Offering mentorship in LLM models, GEN AI, GPT systems, MLOps, data warehousing, and innovative system design. Let's propel your career to new heights!

Merito, Remote   **(6 months - Sep 2022 - Feb 2023)**
CTO   (Data Science, AWS, Agile Methodologies, Leadership, Team Management, Collaboration)

- Created and executed a comprehensive **technology roadmap to adopt cloud computing** and machine learning technologies, resulting in improved business outcomes.
- Led a team of **7 developers, data scientists, designers, and engineers**, providing guidance and mentorship to ensure successful project delivery.
- Transformed the company's management approach by **shifting** from **click-up to Google Sheets, streamlining planning, development, deployment, and testing**.
- Developed a robust process pipeline for product new feature building, feature testing, technical improvements tracking, coding guidelines for performance and security, new ideas tracking, and bugs and issues tracking.
- Built an **in-house data science platform single-handedly within 10 days**, empowering data scientists to develop and maintain all data science activities.
- Managed the **technology budget, ensuring that resources were allocated efficiently and investments** were made in the right technologies Ensured the security of the company's technology systems, including data privacy, network security, and compliance with relevant regulations.
- Collaborated with other departments, including **marketing, sales, and customer support, to ensure that technology solutions** were aligned with **business goals** and customer needs.
- Successfully implemented **AWS infrastructure for the company**, leading to significant cost savings and improved operational efficiency and enhancing the **product's user interface and experience**, moving it from a basic level to an advanced level.

- o Collaborated with cross-functional teams to design and implement data-driven solutions that improve business processes and customer experience.
- o Created and maintained a live workspace on Shopify, increasing online presence and driving customer engagement - http://bit.ly/3ZA388b

StepSetGo, Remote **(Layoff - 6 months - Jan 2022 - Aug 2022)**
Data Scientist (Python, Sklearn, Ensemble Methods, Pandas, Numpy, SQL, Docker, GCP)
- o The developed Fraud detection system for the user on the application.
- o Deactivated approx. 0.91% fraud users from applications.
- o Create a machine learning pipeline using apache airflow for fraud detection.
- o Deployed the project on the GCP cloud using docker.

KEY PROJECTS

PAT Cycling Fraud Detection **(April 2022 - August 2022)**
Key Skills: (Python, Clustering Algorithms, BigQuery, GCP, Docker, Classification Algorithms)
- o In the SSG app on the mobile app as well as smart watches, there are some users who generate some fraudulent cycling activity to take advantage of the prizes and rewards.
- o So we analyzed the mobile as well as IoT data and developed a classification model to detect fraud user/s at the SSG app.
- o Also, detect for which fraud reason/s the user marked or detect as a fraud user.
- o **Deployed it to the server**

Steps Fraud Detection **(Jan 2022 - March 2022)**
Key Skills: (Python, Clustering Algorithms, BigQuery, GCP, Docker, Classification Algorithms)
- o In the SSG app on the mobile app as well as smart watches, there are some users who generate some fraud steps to take advantage of the prizes and rewards.
- o So we analyzed the mobile as well as IoT data and developed a classification model to detect fraud user/s at the SSG app.
- o Also, detect for which fraud reason/s the user marked or detect as fraud user.
- o **Deployed it to the server**

EasyRewardz Software Service, Remote **(1.9 yr - April. 2020 - Dec 2021)**
Data Scientist (Python, Statsmodels, Spark-Streaming, Storm, Event Hub, Kafka, Numpy, SQL, Azure)
- o Improved **data mining processes**, resulting in a 20% decrease in time needed to infer **insights from customer data** used to **develop marketing and finance strategies**.
- o Used **predictive analytics** such as machine learning and data mining techniques to forecast company sales of new products with a 95% accuracy rate and perform A/B testing.
- o **Developed ETL** for data sources used for reporting by **sales, inventory, advertising, and marketing departments**.
- o **Design Wireframes** of the products and **deployment architecture** designing the products and their deployment.
- o **Build**, and **maintain** tools and **infrastructure** for data processing for AI/ML development initiatives
- o Manage ETL and machine learning model lifecycle: **develop, deploy, monitor, maintain**, and **update data** and **models** in **production.**

KEY PROJECTS

NLP Whatsapp Retail Chatbot **(July 2021 - Nov 2021)**
Key Skills: (NLP, Python, Universal encoder, Redshift, Flask, Whatsapp API, Intent Classification)
- o Build an **NLP-powered adaptive Chabot** for brands (like Bata, Pantaloons etc.) to tackle **customer queries/interests smartly**.
- o It needs to **identify intents**, **sub-intents**, and entities from the **user input** queries.
- o Indicates the exact motivation behind customer queries. The predefined intents are – **Loyalty**, **Store**, **Product**, **Greetings**, **Profanity**, **Confused**, **Invoice**, What's New, Shoe Size, and Unknown.
- o Additionally, this chatbot also incorporates attractive features like - **Auto-Correction, Auto-Suggestion** & **Profanity Filtering**.
- o **Deployed it to the server**

Joey Chatbot (Text to SQL) **(March 2021 - June 2021)**
Key Skills: (NLP, Scala, Python, Universal encoder, DNN, Flask, Elastic Search, Docker)
- o NLP Based **chatbot** for the company website using **Universal encoder and cosine similarity.**
- o Take user query then convert into SQL query and give results to users.
- o Also, give the functionality to search using voice commands as google assistant and Siri have.
- o **Deployed it to the server**

CorrectMe (An AI-based text editor) **(Dec 2020 - Jan 2021)**
Key Skills: (Python, Python Flask, Torch, Transforms, JS, HTML)
- o An AI-based text editor that **auto-corrects**, **auto-suggests**, **corrects grammar mistakes** and detects **abusive words**.
- o Used **Bert algorithm** for auto-suggestion, and text blob for auto-correction.

Recommendation System   **(Sep 2020 - Sep 2020)**
Key Skills: (Pyspark, Pyspark.Sql, Pyspark.MLlib, Numpy, Scikit-Learn, ALS, Collaborative Filtering)
o   Recommended **new category and subcategory** using **ALS** based on their previous transaction history.
o   It analyzes the purchase pattern of multiple customers in the **same item, multiple items purchased by the same customers** and the relations between those two aforementioned.
o   Three features involved **user ID, item ID, and quantity**. Data is prepared after **simple aggregation** over customer mobiles.
o   **Deployed it to the server**.

System login using Face Recognition   **(July 2020 - July 2020)**
Key Skills: (Python, PHP and Machine Learning Library OpenCV)
o   This application **detects human faces** in images, offline or in **real time**.
o   We take an **image** as input and process it with **OpenCV Framework** and give output with **human face detection** and other parameters like the name and if that **name exists** in DB then it permits us to **log in to a system** otherwise **denied**.
o   **Deployed it to the server**.

Customer Churn Prediction   **(Aug 2020 - Aug 202)**
Key Skills: (Python, Classification algorithms, Sklearn, Pandas, Numpy)
o   Have customer data with **features,** customer detail, shopping details, etc and with **label churn or not churn**.
o   Train customer data on the **Classification algorithms** like **Logistic Regression, SVM**, Decision Tree, Random Forest and **Neural Network**. Then **predict whether the** company will **churn or not**.
o   **Deployed it to the server**.

Extra work
UI designing wireframes and decisions of the products.
Resolving issues at the production.
Distributed training and/or GPU-based training.
Deployment architecture designing of the products and then their deployment on our test server for testing purposes.

BridgeLabz Solutions LLP. (Mumbai Branch, MH)   (**1 yr 11 months -** June. 2018 - March 2020)
Machine Learning Engineer   (Python, BigQuery, Spark-Streaming, Storm, Event Hub, Kafka, Pytorch, Pandas)
o   **Selecting features**, building and **optimizing classifiers** using machine learning techniques.
o   Hands-on experience in Machine Learning algorithms like **Linear Regression, Logistic Regression, SVM, Decision Tree, Random Forest, Neural Networks, LSTM, NLP & Computer vision.**
o   **Processing**, **cleansing**, and **verifying real-time data** integrity of data used for analysis.
o   **Mentorship** to fellowship engineers on Machine learning and **skill up** them on **future technologies**.

Key Projects
BridgeLabz Website Chatbot   **(Mar. 2019 - Nov 2019)**
Key Skills: (NLP, Python, Sequence Model, Keras, Flask and MongoDB)
o   NLP Based **chatbot** for the company website using **Random forest**, and **SVM** algorithms.
o   Help **Fellowship engineers** to **answer** simple questions. It makes things **easy** and **time-saving** for the company.

Sales Lead Classification   **(Nov 2018 - Apr 2019)**
Key Skills:(Python, Classification algorithms, Sklearn, Keras, Flask and MongoDB)
o   Have sales data with **features** company **size, requirements, name**, location, funded, etc and with **label qualified or unqualified**.
o   Train sales data on the **Classification algorithms** like **Logistic Regression, SVM**, Decision Tree, Random Forest and **Neural Network**.
o   Then **predict** whether the company will be **qualified or unqualified**.

Hidden Talent R&D Pvt Ltd. (Bhilwara, RJ)   (**1 yr -** June. 2017 - May. 2018)
Jr. Software Developer   (HTML, CSS, JavaScript, Material Design, Angular6, JSON)
o   Bringing a designer's concept to life with **HTML, CSS, and JavaScript**.
o   Production, modification, and maintenance of websites and web application user interfaces.
o   Contributing some **back-end experience**, collaborating on **APIs**, and more.
o   Maintaining software workflow management with a project management tool like **GitHub** and **task runners like bitbucket**

Key Projects
Validation Chrome Extension   (Angular6, Material Design, Local Storage)
o   Validation assistant is a **google chrome plugin** and validation helper tool.
o   Using a validation **Assistant we can share test users** with our team. Mainly this **tool is used** by the **Union Bank of America**.

System Logs Management   (Angular6, Material Design, Local Storage, NodeJs, Python)

- o Developed a **python script** to **filter the logs** based on **keywords** and then developed a **Nodejs API** to execute and read the script.
- o Developed UI to showcase the filter logs & **integrate the API** with Angular 4+. Mainly this **tool is used by the Union Bank of America**.

PROBLEM SOLVED

Text2kode   (NLP, Python, Rasa ChatBot, Flask and Bootstrap)
- o Solved company co-workers' problems using **Rasa Chatbot. Searching commands** on google and **getting links** instead of commands.
- o Built an **NLP-based search engine** where users could search for commands and get them in **one line of a word**. **Watch Now**

Amie - Phone Advisor   (Facebook Messenger, Shell Scripting, PHP, MySQL and NLP)
- o Solved **phone buying problem** using an **AI-based chatbot i.e. Amie**
- o They will **ask** the **user** their **requirements** and according to that, it will give some **solutions** to give a clear picture.
- o It **saves a lot of time** that users waste while **searching** and also they will get **more accurate results**. **Watch Now**

EDUCATION

- o **Bachelors** in Information Technology| the University of Rajasthan at MLVTEC| Bhilwara, RAJ. | **Pct: 60**   **Completed:** June 2017
- o **SSC** in Science and Mathematics | CBSE at St Anselm's Sr. Sec. School | Bhilwara, RAJ. | **Pct: 68**   **Completed:** June 2013

ACADEMIC PROJECTS

Image-Based Object Recognition Web App (team major project)   (Bootstrap, Python, PHP and Machine Learning Library OpenCV)
- o This application **detects human faces** in **images** as well as in offline or live **videos**.
- o We take an **image** as input, process it with OpenCV Framework, and give output with **human face detection** and other parameters like the name that detects a human face, its **gender**, its **age**, etc.

JanHeet Android App (team minor project)   (Android SDK, Java, PHP, MySQL, JSON, Restful API)
- o This application fills the **communication gap** between **ward members & councillors** & is also supported by **Swachh Bharat Abhiyaan**.
- o From this, the user can post their problems by clicking the photo of that problem & directly sending it to the Councillor of that ward.

CERTIFICATIONS

- o **Machine Learning** by **Stanford University** certification from Coursera by Professor **Andrew Ng Certificate**
- o **Neural Networks and Deep Learning** certification from Coursera by Professor **Andrew Ng Certificate**
- o **Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization** certification from Coursera by Professor **Andrew Ng Certificate**

ACHIEVEMENTS AND EXTRACURRICULAR ACTIVITIES

- o Secured **3rd position in Ethical Hacking competition** at all India level, organized by NNSC and E-Cell in IIT Mumbai.
- o Founder of a club known as the **Learning Club** consisting of >60 members
- o Delivering **lectures on Web Development** in college, for the students of the 2nd and 3rd year.