Avinash S. Bhad

Data Scientist

Strategic Data Scientist with proven success in designing data-driven solutions for complex challenges. Expertise in Computer Vision, Image Processing, and Financial domain analysis. Proficient in developing and deploying machine learning models (including deep learning) for low-power edge devices. Skilled in statistical modeling, time-series forecasting, and natural language processing (NLP). Eager to collaborate with cross-functional teams to optimize processes and deliver impactful results in a reputed organization

Contact

- **Address** : Flat No. B-09, Samarth Residency, near Laxmi Mata Mandir, Balewadi,Pune - 411045.
- **Ph : 8600021809**
- **Email**: avi.bhad92@gmail.com

Education

B.Tech | B.E Mechanical | 2017 |
J.S.P.M, Pune |
S.P.Pune University | 64.13

HSC | 2010 |
K.T.H.M College, Nashik |
S.P.Pune University | 66.17

SSC | 2008 |
St.Francis High School, Nashik |
Maharashtra State Board |82.30

Personal Details

Dob: 18-09-1992
Hobbies: Drawing, Painting, Photography, Strategy and Puzzle games, Driving , Trekking
Marital Status: Married
Languages: English | Hindi | Marathi

Technical Skills

- Machine Learning ,AI and Data Science
- **Python/ML Packages**: NumPy, Pandas, Sci-Py, Scikit-learn, Seaborn, Matplotlib, Flask.
- **Machine learning**: Linear Regression, Ridge & Lasso, Logistic Regression, Naïve Bayes lassifier, KNN, SVM, Decision Tree, Random Forest

Experience

March 2021 to Present | 3.3 yrs.

Xoriant Pune

Data Scientist

A results-driven Data Scientist with proven expertise in Python, machine learning, and diverse domains including finance and document classification. Demonstrated success in:

- **Strategic Problem-Solving:** Translates complex business challenges into actionable data science projects, meticulously analyzing datasets to uncover patterns and inform strategic decision-making.
- **Solution Development:** Designs, builds, and deploys robust machine learning and deep learning models for applications such as churn forecasting, real-time driver attention monitoring, and spam classification.
- **Technical Proficiency:** Proficient in the Python data science ecosystem (NumPy, Pandas, Scikit-learn, TensorFlow, etc.), deep learning frameworks, and NLP techniques. Excels in object-oriented development and version control (Git/GitHub).
- **Data-Driven Optimization:** Skilled in probability, statistics (including Bayesian methods), time-series analysis, and regularization techniques (Lasso, Ridge), with experience optimizing deep learning models for edge devices.
- **Communication & Collaboration:** Effectively conveys complex findings to stakeholders and collaborates with cross-functional teams to identify challenges, brainstorm solutions, and achieve optimal outcomes.

Key Accomplishments:

- Spearheaded multiple data science projects across diverse domains, demonstrating adaptability and a wide range of skills.
- Improved product features and decision-making processes by translating data insights into actionable recommendations.

Feb 2018 - Sept 2020 | 2 yrs.

Digifleet Transport and Services Pvt Ltd, Nashik

M.D

1. Data-Driven Decision Making:

- **Strategic Planning:** Lead the development of a data-driven business strategy. This involves identifying key performance indicators (KPIs) across operations, logistics, and customer service. You'd then translate these into actionable goals using data insights.
- **Investment Decisions:** Utilize data analysis to evaluate the effectiveness of existing operations, identify areas for improvement, and make informed decisions about resource allocation and technology investments (e.g., fleet management software, predictive maintenance).

2. Operational Optimization:

- **Fleet Management:** Leverage data from GPS tracking, sensors, and route optimization software to improve fleet efficiency, reduce fuel consumption, and optimize delivery schedules.
- **Customer Service Enhancement:** Analyze customer data to understand their needs and preferences. Use these insights to improve customer service offerings and personalize interactions.

3. Risk Management and Compliance:

- **Safety and Security:** Implement data-driven safety programs to monitor driver behavior, reduce accidents, and ensure regulatory compliance.
- **Fraud Detection:** Leverage data analysis to proactively identify and prevent fraud in deliveries, payments, and resource usage.

Ada-Boost, Gradient Boosting, XGBoost, K-means Clustering.
· **Text Processing**: NLTK, Term Frequency-Inverse Document Frequency (TF-IDF), Word2Vec, Bag of Words. doc2vec, sent2vec, keypharse extraction
· **Languages**: Python.
· **Cloud Platforms/Services**: AWS. EC2, Sagemaker, Notebook instance, AWS container, S3, Multi-cloud environment, Azure, Deployment
· **Time Series Analysis**, AR,MA,ARMA,ARIMA,SARIMA
· Deep Learning Frameworks: TensorFlow, Keras, PyTorch
· Big Data Technologies: Spark, Hadoop, Kafka
· Data Visualization: Matplotlib,Seaborn, Plotly, Bokeh, or D3.js
· Databases: PostgreSQL, MySQL
· Version Control: Git, GitHub
· Machine Learning/AI Specialties
· Computer Vision: OpenCV.
· NLP: spaCy, Gensim, Transformers (like BERT)
· Reinforcement Learning: Stable-Baselines, Gym
· Cloud & Deployment

Others Skills

•**Web stack**: Flask, Postman
•**Operating Systems**: Linux, Windows.
•**Database**: SQLite, MongoDB
•**Deep Learning**: Neural Networks, ANN, CNN, DNN, Transfer Learning, Back Propagation, Tensorflow 2.x, Keras
•**Math's & Stats**: Filter, Wrapper, Embedded Methods, P-Value, T-Test, Z-Test, ANNOVA test, Chi-Square Test, Info-Gain Test, Hypothesis Testing. Probability, statistics, linear algebra, probability, statistics, linear algebra, Gradient Descent
•**JIRA**: Tickets, Tasks, Reports.
•**Methodology**: Agile
•**GitHub**: Git, Github, Gitbash
•**Anaconda**: Conda Env Setup.
·IDE: Jupyter Notebook code.
·Fix bugs in code.

Project Details

Project 1 : Churn Forecasting and Analysis (MAchine Learning)

Description:

- Developed a predictive model to identify customers at high risk of churn within the telecom industry.
- Performed extensive data cleaning and preprocessing of customer demographics, usage patterns, and support interactions using analytical abilities.
- Explored various machine learning algorithms (e.g., Logistic Regression, Random Forest, XGBoost) and evaluated their performance using relevant metrics (accuracy, precision, recall, F1-score).
- Implemented feature engineering techniques to enhance model performance and identify key churn drivers.
- Presented findings to stakeholders, translating complex data insights into actionable recommendations for customer retention strategies using innovative solutions.

Roles & Responsibilities:

- **Data Exploration & Preprocessing:** Cleaned, transformed, and visualized data to uncover trends and potential predictors.
- **Feature Engineering:** Created and selected informative features to improve model accuracy.
- **Model Development & Evaluation:** Built, trained, and compared the performance of various machine learning models.
- **Hyperparameter Tuning:** Optimized model parameters for the best performance.
- **Results Interpretation & Communication:** Presented insights and recommendations to stakeholders in a clear and actionable manner.
- **Benifits to the client** : Increased Market Share,Enhanced Brand Reputation,Improved Customer Satisfaction

---

Project 2 : Driver Attention Monitoring System (Deep Learning)

Description:

- Developed a real-time driver distraction detection system using deep learning and computer vision techniques.
- Collected and annotated a diverse image dataset of drivers to train convolutional neural network (CNN) models.
- Implemented image preprocessing techniques (resizing, normalization, augmentation) to improve model robustness.
- Experimented with various CNN architectures (e.g., VGG, ResNet) and optimized for both performance and edge deployment.
- Developed visualization techniques to interpret model decisions and understand the focus areas for distraction classification.

Roles & Responsibilities:

- **Dataset Preparation**: Collected, labeled, and preprocessed image data for model training.
- **Model Architecture**: Designed, built, and evaluated different CNN architectures.
- **Transfer Learning**: Leveraged pre-trained models (if applicable) and fine-tuned them for the driver attention task.
- **Optimization & Deployment**: Optimized the model's size and inference speed for real-time performance, potentially targeting edge devices.
- **Visualization & Analysis**: Interpreted model outcomes to understand key features indicative of driver distraction.

Project 3 : NLP based Segmentation Protocol for Predicting Diseases

Description:

- Developed an NLP system to analyze medical text and predict diseases.
- Preprocessed medical data (doctor's notes, etc.) using text cleaning and normalization.
- Applied NLP techniques (NER, topic modeling) to extract key symptoms and medical terms.
- Built predictive models using machine learning or deep learning techniques.
- Evaluated and refined model performance using relevant NLP metrics.

Roles & Responsibilities:

- **Data Preparation:** Cleaned and prepared medical text data for analysis.
- **NLP Implementation:** Applied NLP techniques for feature engineering and entity recognition.
- **Model Development:** Built and evaluated machine learning/deep learning models.
- **Optimization & Refinement:** Improved model performance through iterative development.
- **Benifits to client**: Cost Savings,Faster and More Accurate Diagnoses,Improved Data Collection and Analysis,Reduced Physician Workload