Abhay Kumar

Bangalore,Karnataka,India

Email : **abhay8280@gmail.com** | Phone : **+91-9731-812-812**

LinkedIn Profile | Github Profile

About Me:

Passionate and accomplished Machine Learning Professional with a decade of hands-on expertise in Computer Vision, Natural Language Processing, Recommender Systems, and complex problem-solving. Proven track record of leading teams, engaging with clients, and contributing impactful research to the field. A dedicated individual with a passion for innovation and a drive to deliver cutting-edge solutions.

Objective

With my extensive experience, I am well-equipped and prepared to take on roles, in the field of Machine Learning, where I can drive impactful innovation, mentor talent, and contribute to organizational excellence and growth.

Professional Experience:

Machine Learning Lead - Target Corporation (Feb 2021 – Present)

Anomaly Detection System for Data Pipeline

- Developed an anomaly detection system to prevent corrupt images from entering the data pipeline.
- Identified and blocked 30% of anomalous images, resulting in substantial reductions in processing time
- Improved downstream algorithms, leading to an 8% enhancement in business metric.

Invoice Digitisation System

- Implemented OCR model fine-tuning to detect specific fonts used in invoices.
- Conducted document analysis to locate table areas for text extraction.
- Identified and reconstructed tabular structures.
- Digitized table data into structured JSON format, preserving row and column information as keys.
- Extracted key attributes from invoices for image indexing and efficient searching.

  Significance:

- Streamlined the digitization process, reducing manual effort significantly.
- Improved data accuracy and accessibility for efficient search and retrieval.
- Enhanced overall document management and data extraction efficiency.

Real-Time Safety Monitoring System

- Developed a real-time monitoring system using camera streams to detect individuals in unsafe zones and ensure adherence to safety protocols.
- Recorded safety violations with evidence, raising safety incidents for review and training.
- Extended monitoring capabilities to cover a wider area, overcoming resource and weather constraints.
- Enhanced safety and operational efficiency.
- Reduced safety violation events by 40%.

Privacy Blur Feature Implementation

- Developed a privacy blur feature for archived video streams.
  - Employed lightweight object tracking and detectors to reduce processing needs.
  - Ensured real-time video processing while obfuscating individuals' identities.
  - Improved security and privacy standards of the video archive system.
  - Collaborated for seamless integration and performance optimization.
- Mentored junior team members, fostering their growth and enhancing the overall team's performance.

Senior Data Scientist - Embibe (May 2019 – Jan 2021)

Question Tagging and Topic Extraction System

- Designed and implemented an automated system to tag questions with subjects and extract relevant topics from text.
- Enhanced content organization and searchability in a repository of millions of question papers.
- Utilized natural language processing techniques for precise content discovery.
- Developed question tagging using both text and image content, enabling better content recommendation and search accuracy.

In-House Annotation Tool Design

- Spearheaded the development of a custom in-house annotation tool, incorporating user and project management workflows.
- Collaborated closely with Subject-Matter Experts (SME) and annotators to gather requirements, ensuring that user needs were at the forefront of the tool's design.
- Integrated active learning capabilities into the system, enabling it to learn from initial observations evolve the annotation flow into an audit flow, where system-generated annotations were verified by annotators, leading to the rapid annotation of thousands of images with a small annotator team.

Senior Data Scientist - ClipIndia (Sep 2018 – Apr 2019)

- Implemented NSFW video classification for Content Moderation, maintaining a clean platform by automatically detecting obscene content, leading to a 28% reduction in inappropriate content.
- Designed a video asset tagging system with multiple attributes, improving content discoverability and feed ranking, resulting in a 3% increase in DAU.

Senior Data Scientist - Cogknit Semantics (Sep 2015 – Sep 2018)

catalogue filtration project

- Spearheaded the catalogue filtration project, ensuring accurate product categorization on the .com website, resulting in a 12% reduction in incorrect product placements.
- Collaborated directly with clients to understand their specific business requirements and challenges that could be addressed through our Machine Learning capabilities.
- Led a high-performing team, taking responsibility for collecting client requirements, translating them into technical specifications, and effectively communicating these requirements to the team.
- Played a key role in project planning and estimation, ensuring alignment with client expectations, and optimizing resource allocation.
- Orchestrated end-to-end solution development, including data collection, preprocessing, model development, testing, and deployment, to meet client objectives and project timelines.
- Successfully delivered customized Machine Learning solutions, exceeding client expectations and achieving project milestones.

course recommender system

- Led the development and implementation of a course recommendation system within our Learning & Development (L&D) platform.
- Collaborated with cross-functional teams to gather user data, including user behavior, preferences, and course interactions.
- Utilized machine learning algorithms, such as collaborative filtering and content-based filtering, to personalize course recommendations for users.
- Integrated data pipelines for real-time data ingestion, processing, and updating of recommendations.

Sr. Research Analyst - Edureka (Dec 2013 – Sep 2015)

- Conducted extensive research on Machine Learning course content and collaborated closely with Subject Matter Experts (SMEs) to develop comprehensive technical content, including exercises, quizzes, and other learning materials.
- Served as a Teaching Assistant (TA), providing invaluable support to students in solving complex technical problems and facilitating their learning experience.
    - Recognized as the "Best Performer" for five consecutive months for outstanding contributions to student success and engagement.
- Contributed significantly to the community by authoring blog posts on advanced Machine Learning and Big Data topics.
    - Five of my blog posts became the highest-traffic generators on the platform.
- Led the initiative to create a comprehensive knowledge base, reducing the workload on the TA team and enhancing the overall efficiency of support and learning resources.

Education:

- Bachelor in Computer Science, Visvesvaraya Technological University (VTU), 2013.

Skills:

- Computer Vision: Object Detection, Image Classification.
- NLP: Text Classification, Topic Modeling.
- Machine Learning: Deep Learning (PyTorch, TensorFlow), Scikit-learn.
- Data Processing: NumPy, Pandas, Spark, Hadoop.
- Cloud Platforms: AWS, Azure, Google Cloud.
- Software Engineering: Git, Docker, CI/CD.

Publications and Presentations:

- Poster presentation on "Domain Adaptation of Image Caption Model for Video Descriptions" at AI & Deep Learning Conference (GTC 2018, NVIDIA).

References:

Available upon request.