# Abhay Kumar

Bangalore, Karnataka, India  
Email: abhay8280@gmail.com | Phone: +91-9731-812-812  
[LinkedIn Profile](https://in.linkedin.com/in/awesomeabhay) | [Github Profile](https://github.com/abhaymise/abhaymise.github.io)

## About Me

Experienced Machine Learning Professional with expertise in Generative AI,LLM, VLM, NLP, Computer Vision and solving complex problems.

## Professional Experience

### Sr AI Engineer - Target Corporation (*Feb 2021 – Present*)

#### Invoice Information Extraction System

* Built a ML-based system using OCR,LayoutLM,Table Detection and LLMs to automate invoice data extraction, reducing manual effort.
* Integrated the system with the company’s invoice processing platform, enabling seamless automation of invoice data extraction and processing.

#### Nutrient Information Extraction from Food Labels

* Developed a similar system for food labels, ensuring accurate nutritional tracking and seamless integration with metadata management.

#### RAG powered Document Query System

* Implemented a retrieval-augmented generation system for querying large document corpora and summarizing key insights.
* Integrated the system with the company’s document management platform, enabling users to search for information across a wide range of documents.

#### Anomaly Detection System for Image Data Pipeline

* Built a system to identify outliers like corrupted images and inconsistencies.
* Integrated with data pipelines for continuous monitoring and quality control.

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

#### Question Tagging and Topic Extraction System Development

* Applied NLP techniques for automated topic assignment and question tagging
* Integrated the system with the company’s question bank platform improving content organization & discovery

#### Query Intent Classification and Routing System

* Built an intent classification system to optimize query routing, enhancing response times and customer support efficiency.

#### Data Annotation Tool Design

Designed in-house annotation tools with active learning, enabling rapid, high-quality data labeling at scale.

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

#### NSFW Video Classification

* Developed a system to detect and filter inappropriate content, improving platform safety.

#### Video Asset Tagging System

* Built a tagging system that boosted daily active users (DAU) by 3% through improved content discovery.

#### Review Moderation System

* Implemented a system to detect and filter NSFW comments with 95% accuracy, enhancing user trust.

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

#### Catalogue Misplacement Detection System

* Developed a CV-based system to detect and correct displacements in product catalogues.
* Led to cleaner & accurate product catalogues, reducing incorrect placements by 12%.

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

* Researched ML content, collaborated with SMEs, authored technical blogs, and built a knowledge base to improve learning resources and reduce support workload.

## Education

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

## Skills

* **Generative AI**: LLMs, RAG, LangChain, Agents, ChatGPT, BERT, LLAMA
* **Search**: Semantic & keyword search, re-ranking, Elasticsearch, FAISS, Milvus
* **Computer Vision**: Object detection, classification, segmentation
* **NLP**: Text & intent classification, topic modeling
* **ML & Data**: PyTorch, TensorFlow, Scikit-learn, Hugging Face, Pandas, Spark
* **Cloud & MLOps**: AWS, Azure, GCP, Docker, CI/CD, MLflow, APIs

## Publications and Presentations

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