Project 1: Entity Extraction Module for Insurance Client

Duration: 14 months (EXL)

Challenge:

Develop an automated system to extract key details (excess amount, deductible, sum insured, claim eligibility) from complex insurance documents (policy & claim schedules) containing multiple claims per number.

Skills Leveraged:

- **Document Processing:** Extracted raw text using tools like pypdf, pymudf, and Amazon Textract.
- **Information Extraction (IE):** Employed advanced techniques like LayoutLM with Detectron for invoice extraction within schedules.
- **Natural Language Processing (NLP):** Utilized large language models (LLMs) like llama2 and FastChat for handling complex amounts. Used chroma db and FAISS as vector store.
- **Entity Recognition & Regular Expressions (RegEx):** Extracted additional details like dates, keywords-based extraction using SpaCy models and RegEx.

Project 2: Entity Extraction, Conversational BI & PPT Generation for Retail Client

Duration: 7 months (EXL)

Challenge:

The client, a retailer, faced challenges managing product data from various brands' catalogues (unstructured PDFs) and generating insights from sales data.

Skills Leveraged:

- **Document Understanding:** Extracted text and images from PDFs using AWS Textract and image parsers.
- **Data Transformation:** Converted unstructured data into a structured format for easy analysis.
- **Conversational AI:** Built a chatbot powered by ChatGPT4 to facilitate communication and provide insights.
- **Business Intelligence (BI):** Enabled data analysis through SQL queries and report generation.

Project 3: Logo Recognition & Analytics for Sports Broadcasts (Entertainment Client)

Duration: 2.5 months (EXL)

Challenge:

The client, an entertainment company, needed a system to automatically track sponsor logos appearing during televised sports broadcasts. This information is crucial for accurate billing based on logo exposure time and size.

Skills Leveraged:

- **Logo Detection:**
  - YOLO (You Only Look Once) object detection algorithm was employed to identify and localize sponsor logos within video frames.
- **Data Extraction:**
  - AWS Textract (potentially) assisted in extracting additional text data relevant to logos (optional, depending on specific implementation).
- **Annotation Tool:**
  - An annotation tool likely facilitated the training process for the logo detection model by providing labeled data sets containing sponsor logos.
- **Data Storage & Analytics:**

    - A SQL database stored information on detected logos, including:
        - Sponsor name
        - Appearance count
        - Size of logo appearance

Project 4: Fraud detection using claim images and information extraction

Duration: Ongoing (EXL)

Challenge:

The client is an insurance company, needed a system to automatically detect images of cars and property damage submitted by clients to process claims. This information is crucial for accurate disbursement of insurance amount.

Skills Leveraged:

- **Image similarity search:**
    - Leverage multiple models (Open AI pre-trained CLIP, Resenet50, DeepImageSearch, VIT, VGG) for image feature extraction and similarity search. Used Chroma DB for image embedding vector storage.
- **Data Storage & Analytics:**
    - A SQL database stored information on detected image information, including:
        - Similar image score and similar images which can be potentially photoshop.
        - Claimnumber

Project 5: Streamlined Insurance Decision-Making for Transportation Client

Duration: 8 months (Coforge)

Challenge:

The client, a transportation company, faced a time-consuming process of evaluating insurance proposals from brokers. Underwriters had to manually extract key details from emails containing attachments (policy documents, spreadsheets) and email bodies.

Skills Leveraged:

- **Automated Data Extraction:**
    - Microsoft Graph API retrieved relevant email data.
    - A BERT model, trained on underwriter-provided keywords, extracted crucial entities from email bodies and attachments (PDFs, Excel sheets) processed by pypdf, pdfplumber, and pymudf.
- **Centralized Information View:**
    - The plugin consolidated all extracted information onto a single web page, eliminating the need to access individual attachments.

Project 6: Marketing Campaign Analytics & Optimization for Retail Client

Duration: 28 months (TCS)

Challenge:

The client, a retailer, needed to optimize their marketing campaigns across email, mail, and calling channels to improve campaign performance and ROI.

Solution: Analyzed past marketing campaign data (email, mail, call) to build models (Logistic Regression, Gradient Boosting) for predicting performance. Used these

models to optimize future campaigns and maximize ROI through data-driven insights.

Achievements:

Consecutive Stellar Performer: Recognized as a top performer at EXL Services for two consecutive years (2022, 2023) due to exceptional contributions in designing, building, and deploying successful AI projects for major clients

On the spot awards: Awarded multiple times at TCS (2016, 2017) for swiftly resolving critical issues, demonstrating strong problem-solving abilities.