

	<p>An AI-based pipeline that extracts key fields from a wide variety of insurance forms for downstream processing and claims automation.</p> <p>Key Contributions:</p> <ul style="list-style-type: none">● Engineered OCR-first document parsing pipelines using Tesseract, EasyOCR, and fallback image preprocessing techniques.● Used PDFMiner and OpenCV to isolate structured zones such as checkboxes, signature fields, tables, and input blocks.● Applied document template matching to route input documents to appropriate parsing logic, increasing template match rate by 30%.● Developed a hybrid approach using rule-based parsing with prompt-based LLMs (like Mistral and LayoutLM) for improved field recognition.● Containerized services using Docker, and integrated them with Kafka-based message queues for scalable ingestion.● Enabled logging, retry mechanisms, and failure tracking for documents with low confidence scores or extraction errors.● Built validation dashboards to allow manual correction and learning feedback loop to further refine extraction logic.● applications.● Implemented logging and monitoring tools for proactive issue resolution.● Provided technical support and mentorship to junior developers.
--	--

Project Title:	Certificial
Description:	Insurance form extraction service where different types of insurance forms details has to be extracted for further downstream systems
Responsibilities	<p>Collaborated closely with business analysts to gather and refine specifications, ensuring alignment with evolving business needs and accurately addressing defects or enhancement requests.</p> <p>Participated actively in the entire software development life cycle, including coding, debugging, performance optimization, and production support.</p> <p>Conducted high-level design and requirement elicitation sessions to translate business processes into technical specifications.</p> <p>Engineered a document intelligence system capable of accurately extracting</p>

	<p>An AI-based pipeline that extracts key fields from a wide variety of insurance forms for downstream processing and claims automation.</p> <p>Key Contributions:</p> <ul style="list-style-type: none">● Engineered OCR-first document parsing pipelines using Tesseract, EasyOCR, and fallback image preprocessing techniques.● Used PDFMiner and OpenCV to isolate structured zones such as checkboxes, signature fields, tables, and input blocks.● Applied document template matching to route input documents to appropriate parsing logic, increasing template match rate by 30%.● Developed a hybrid approach using rule-based parsing with prompt-based LLMs (like Mistral and LayoutLM) for improved field recognition.● Containerized services using Docker, and integrated them with Kafka-based message queues for scalable ingestion.● Enabled logging, retry mechanisms, and failure tracking for documents with low confidence scores or extraction errors.● Built validation dashboards to allow manual correction and learning feedback loop to further refine extraction logic.● applications.● Implemented logging and monitoring tools for proactive issue resolution.● Provided technical support and mentorship to junior developers.
--	--

Project Title:	Certificial
Description:	Insurance form extraction service where different types of insurance forms details has to be extracted for further downstream systems
Responsibilities	<p>Collaborated closely with business analysts to gather and refine specifications, ensuring alignment with evolving business needs and accurately addressing defects or enhancement requests.</p> <p>Participated actively in the entire software development life cycle, including coding, debugging, performance optimization, and production support.</p> <p>Conducted high-level design and requirement elicitation sessions to translate business processes into technical specifications.</p> <p>Engineered a document intelligence system capable of accurately extracting</p>