**Selection Criteria Responses**

**Essential Criteria**

**1. Experience building and maintaining cloud services such as Amazon Web Services (AWS) Textract and S3**

* **Sasidhar got certified in AWS Machine Learning Specialisation and extremely comfortable with most of the AWS related services in building a Machine Learning Solution.**
* **Sasidhar has used S3 and Textract in multiple projects. He has built and maintained machine learning solutions for entity extraction from the invoices and medical documents received from the medical providers while he was at ICare for the documents and emails that land into an S3 bucket. The solution aaaa**
* **At Suncorp, Sasidhar also uses the Textract API to extract the text from emails and their attachments to classify the emails and their documents into multiple categories which are then used by the downstream applications. This model is hosted as an API on a platform called Domino Datalabs.**
* **Also, Sasidhar is crucial in experimenting and implementing text extraction from documents using the latest state of the art deep learning techniques using models like LayoutLM.**
* **Sasidhar is also responsible for creating workflows on the Databricks platform that are triggered when a new files lands in the S3 bucket. These workflows are primarily for classifying the incidents into multiple categories.**
* **2. Experience building and maintaining software using Python and Extensible Markup Language (XML).**
* **Sasidhar is a highly experienced data scientist and machine learning engineer. He has been using Python as a primary programming language for building analytical solutions for the past 6 years**
* **He has extensive experience in building and deploying solutions as web services or as batch processes and integrate them with existing enterprise applications. These services are required to follow a standard data format like XML or JSON for communication. Hence he is well versed in using these.**
* **Sasidhar also has extensive experience in building web applications using frameworks like Streamlit using Python for quick development.**
* **Sasidhar also has experience in building distributed applications using Python leveraging the distributed frameworks like Spark or Ray**

**3. Experience with working in Linux and with Application Programming Interfaces (APIs)**

* **Sasidhar, with his rich experience as a software engineer is extremely comfortable in usinh Linux as the primary operating system.**
* **Sasidhar has built cron jobs to schedule the machine learning jobs to run everyday to predict the churn while he was working at Vodafone**
* **Sasidhar also has experience in consuming various APIs , both internal and external to enrich the data.**
* **Sasidhar also has experience in exposing the machine learning services as API endpoints using Flask/FastAPI and AWS Sagemaker endpoints in all the organisations he worked**

**4. Knowledge or experience with unstructured text processing such as regex or entity extraction using natural language processing is highly desirable (bordering on essential)**

* **Sasidhar has extensive experience on NLP as most of his projects deal with unstructured text.**
* **Some of the notable projects that he worked on are :**
  + **Claim Triaging based on the claim details which includes the claim description**
  + **Assessment Prediction based on the claim description..**
  + **Document Classification of emails and its attachments.**
  + **Automated Order extraction from emails.**
  + **Risk Incident classification.**
  + **Data Anonymisation to redact personal information.**
* **All the above projects require entity extraction and language modelling in some form to achieve the desired result.**
* **Sasidhar is quite comfortable using the opensource frameworks like NLTK, Spacy , Gensim that aid In a wide variety of NLP tasks**
* **Sasidhar also has demonstrable experience in using open sourced Large Language models and proprietary models to build solutions to summarise and create chat functionality over enterprise documents.**

**5. A demonstrable passion for innovation, dedication to working collaboratively and the ability to explain or translate complex models and findings to business stakeholders.**

* **Sasidhar always keeps himself abreast with the latest technological advances and brings in them to create innovative solutions.**
* **Sasidhar takes pride in introducing and productionising the zero-shot modelling to classify risk incidents using the open sourced large language models at Suncorp.**
* **Sasidhar also understands the technical and business constraints and builds usable solutions given those constraints , for example , Sasidhar was able to productionise a churn model just using proprietary hardware as there was no machine learning platform available while at Vodafone.**
* **Sasidhar understand that data science is a team game and works collaboratively across the technical and business teams in order to build and productionise models. He has deployed multiple models to production within a short period of time at Suncorp.**
* **Sasidhar has great experience in working with stakeholders by taking them along the journey of solution development and finally presenting the findings and solutions to the business stakeholders. As an example, Sasidhar was able to come up with the custom metrics like Top-N Hit Rate when developing a solution to classify incidents into multiple categories.**

**Desirable criteria  
1. Experience with OCR**

* **Sasidhar has god experience with building document parsing and entity extraction solutions using OCR libraries like Tesseract.**
* **Used the output from the OCR extraction to structure the output as sequential text using the bounding boxes.**
* **Also experimented with multi modal deep learning models like LayoutLM to extract the text.**