Meet Agrawal
Topic - Cloud

Assignment - 1 - PoC Document

Basic Setup -

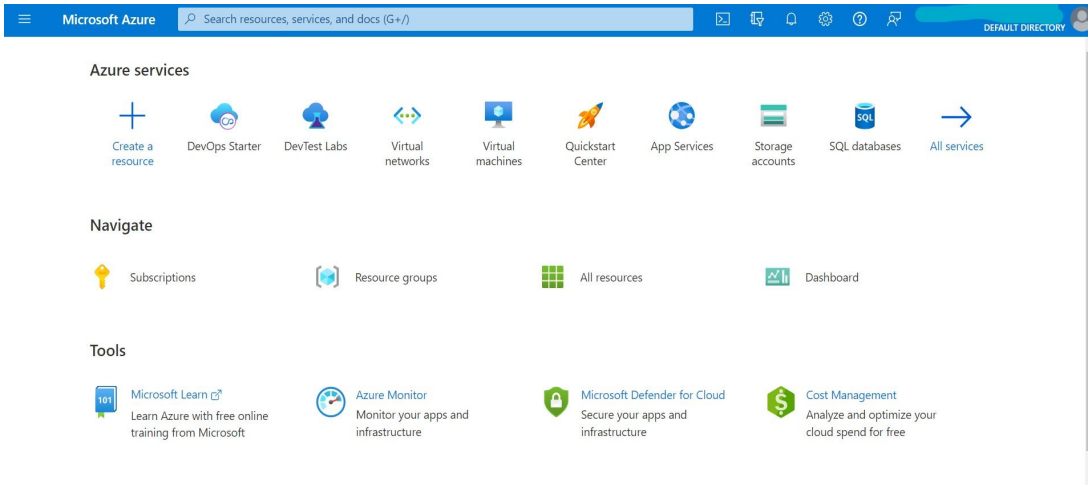
Activating Student Credits for using Azure -

Solve the assignment by activating Azure for Students using your college email id
Use the link - https://azure.microsoft.com/en-in/free/students/

Making account on Azure -

1) Register with a gmail account (not the student account - normal gmail account)
2) Then Click Link above and sign In again
3) You will be redirected to a form where you need to enter your college email
4) Accept confirmation email on your college account and you are all set with the student credits.

Final Screen once done with the setup looks -



Setting up Node JS :

You need Node JS to solve this assignment
Watch the video to setup if you haven't already done -
https://youtu.be/__7eOCxJyow

Steps to follow :

1) Move to your cloud portal on azure

https://portal.azure.com

2) Create a Form Recognizer service

https://ms.portal.azure.com/#create/Microsoft.CognitiveServicesFormRecognizer

3) Create Storage Account

https://docs.microsoft.com/en-us/azure/storage/common/storage-account-create

4) Update Resource Sharing CORS for your Storage Account
https://docs.microsoft.com/en-us/azure/applied-ai-services/form-recognizer/quickstarts/try-v3-form-recognizer-studio#additional-prerequisites-for-custom-projects

5) Within the storage account make a container

6) Container is where you upload your dataset

Link to Dataset -
https://drive.google.com/drive/folders/12Lg5mFofm4qlfCm4FEPdbZg8idwpS2r3?usp=sharing

Dataset Description :

a) It consists of 5 PDF documents which are dummy resumes made using the VNIT resume template
b) A Dataset of size 5 is the minimum requirement for training a model in Azure
c) As the size of the dataset is small and just covers the minimum limit you are free to add more dummy resumes to your container for training the model for better accuracy
d) The file " `1.pdf` " is the resume template itself and this can be used for making up more dummy resumes

7) Use the Form Recognizer studio to label your dataset and run the analysis on some input PDF file

https://fott-2-1.azurewebsites.net/

Reference for labeling dataset using the labeling tool -

https://docs.microsoft.com/en-us/azure/applied-ai-services/form-recognizer/quickstarts/try-sample-label-tool#train-a-custom-form-model

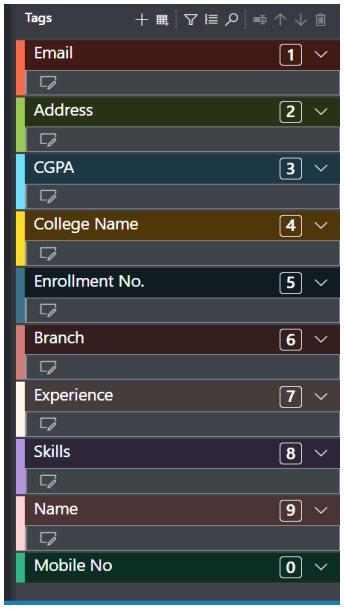
Steps include -

a) CORS setting updated
b) Generating the SAS URL for the storage account
c) Creating a new project on the app link provided
d) Labeling dataset
e) Training the model

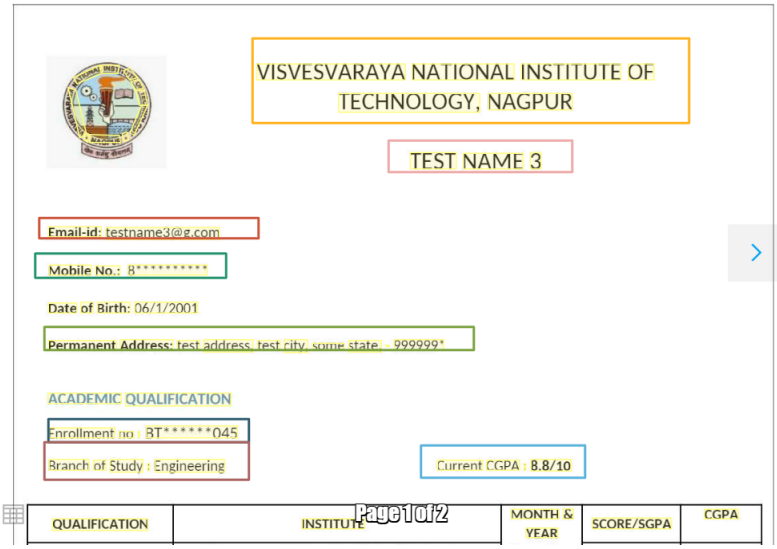
NOTE : Remember to copy the model ID after training
(needed for further reference to the model while testing it)

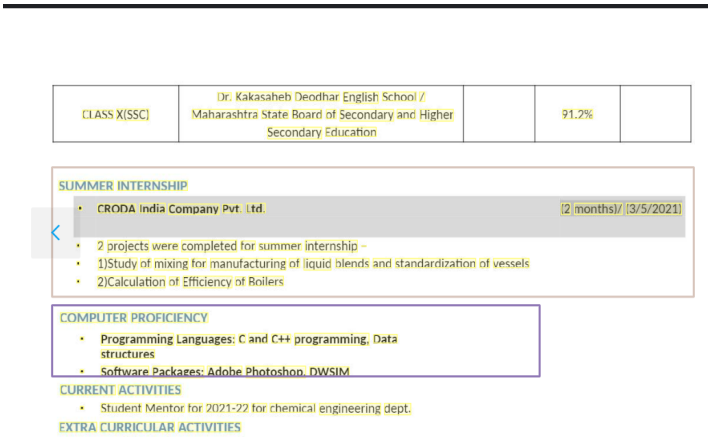
Fields to Label :

a) Name
b) Email
c) Address
d) CGPA
e) College Name
f) Enrollment No.
g) Branch
h) Experience
i) Skills
j) Mobile No.



Sample Labeling :



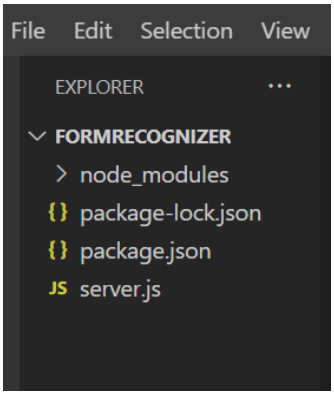


8) Add your own resume for analysis and see the accuracy you achieve from the model

9) Can further improve the training dataset by adding more sample resumes for training the model

File System :

1) Create a FormRecognizer Folder
2) Run " `npm init` " to initialize node in the folder
3) Make a file " `server.js` "
4) Run " `npm i @azure/ai-form-recognizer` " to install the relevant package



10) Look up npm package for "`Recognize Forms Using a Custom Model`"
https://www.npmjs.com/package/@azure/ai-form-recognizer

11) Paste the code in the server.js file and make updates to it as follows

Updates to be made :

```
async function main() {
  const endpoint = "<cognitive services endpoint>";
  const apiKey = "<api key>";
  const modelId = "<model id>";
  const path = "<path to a form document>";
```

Endpoint & ApiKey - To be taken from FormRecognizer Resource in your azure portal

Model-Id - Generated in the tool once you are finished training your model

Path - This is the path to your resume file on your local system

12) Run the file using the command
`node server.js`

13) The analyzed results are printed to the console with their respective confidence values