CC ML RECRUIT TASK PDF

TASK-1

INITIATION: -

Before we begin, I would like to mention that the notebook is built on Kaggle, 1st cell of the notebook is standard Kaggle start, we then reinstall scikit learn and imbalanced learn (for smotenc) to specific versions which are know to work well together. Imported seaborn and matplotlib, and then the dataset is loaded and some info about it is displayed. We then separate categorical and numerical columns. a heatmap of corr matrix is then plotted between numerical features.

FEATURE ESTIMATION: -

We start with feature 1, it has a range of 15-22 as seen from dataset, and has no particularly strong corr with any other column, given such a sweet range, I was inclined to guess it as AGE.

For feature 2, it had a strong corr with grade columns (g1, g2, g3) and neg corr with Dalc and failures, indicating it to be a study related feature, I guessed it to be TIME SPENT STUDYING

For feature 3 , it had a very strong corr with Dalc and goout, which seems to make it a partying feature, I guessed it to be TIME\_SPENT\_ON\_GHAT

LEVEL 2:DATA INTEGRITY AUDIT

We now begin our setup for imputation, we first reload heat map of corr matrix, which will be our basis for imputation of numerical features, as Fedu and Medu had a very strong corr

I simply chose to put Fedu=Medu for missing values.

For other missing numerical features, such as (“traveltime”,”freetime”,” absences”,” G2”,” Feature\_1”,” Feature\_3”), I used linear regression using other columns which with they had high corr to predict their values, I then rounded them off to maintain the data of column.

For categorical columns, I first try to see if we have any other column (both num or cat) which might help us predict, for this I used stripplot and heatmap but sadly there was no clear indication for the same, so I went ahead and imputed randomly between the two values for both column while maintain their current ratio.

OUR DATASET IS NOW READY.

I then check the value\_counts for romantic column , it was imbalanced but not severely,

We now scale our numerical columns using standardscaler, one hot encode the scaled dataset, and also apply pca for knn.

LEVEL 3

We now have to train and check performance of our models , for this task, I decided to use logistic regressor, randomforest,knn,svm,xgboost,catboost,neural-network

As they are quite a lot, I decided to first compare their base performance and then tune our top performer,

I start with logistic regressor, and surprisingly it performed worse when I passed class\_weight=’balanced’ to counter imbalance, max\_iter = 1000 to let the model converge ,

I have used cross-validation to measure accuracy of the models.

Now I have used random forest, with very same procedure as above, I then moved to knn, and again it performed worse on pca dataset compared to ohe dataset ☹. I again used cross validation to measure accuracy. Moving on to SVM, it performed surprisingly good , with kernel set to radial basis , as xgboost expects 1/0 labels instead of yes/no , I used label encoder , XGBclassifier performed bad , although it could have been fine tuned using optuna or Bayesian optimisation, but due to time constraints I could not do that.i then used catboost, a very famous model on Kaggle, with verbose off to prevent spamming on log.

We now move forward onto neural-network

For it I first split the dataset using train\_test\_Split.

I used the Sequential and dense function to create a very basic neural network , with sigmoid function on output for binary classification ( it is the same function used by logistic ) . we compile it and then fit it on our data .

By now , it is our best performing model, with accuracy varying around 0.65 .

But this was predicting poorly for ‘yes’, So

I decide to further tune it by using smotenc ( extension of smote for datasets with both numerical and categorical columns ). I again then implemented smotenc, train\_test\_split our new dataset, and again went through training, we now see a increase in accuracy, with 0.68 , it even went upto 0.73 while executing it repeatedly,

LEVEL 5 SHAP

Here I have imported and applied shap on model 1 (log reg) to explain the importance of features, we then plot a waterfall plot for a student having yes and no output to explain the features which led to the choice. I then implemented decision boundary for this log\_reg.

Next I applied shap for xgboost model , going through the same steps as above.

BONUS  
plot 1 is of random forest due to rectangular boundary lines with high complexity,

Plot 2 is of decision tree due to rectangular boundary with less complexity

*plot 5 is knn due to very adaptive boundary, indicating it has varies heavily with locality*

TASK 2 ML

(sorry, this is going to be short as I am writing this at 1 46 am)

we begin this task by first installing langgraph ,and then upgrading it .we then installed langchain[google\_genai] for our gemini based chatbot.we then imported out chat model, got our api keys, PLEASE ENTER YOUR API KEYS HERE INSTEAD.we defined a very basic chatbot based on google genai gemini 2.0 , which simply returns the reply from this ai.

We add this as our node and them moved on to create a router function . here, if the input contains any math operators, it will redirect it to calculator, if it contains words like trending or fashion it will redirect it to fashion tool , if it contains word weather it will redirect it to weather tool. I then created a calculator tool node which simply evaluate the input PLEASE ENSURE THAT THE INPUT ONLY CONTAINS THE MATH EXPRESSION and returns the answer .for weather node, I used openweather\_api, I first split the message from “weather in” and then used the 2nd part of the split which contains the city name, only use expression of format “\_\_\_\_\_\_\_\_\_ weather in \_\_\_\_\_\_?”.for fashion\_rec, I simply used the gemini model above, for all cases, it just adds “ in context of fashion” to let model know that trends are asked in fashion segment, it then siply returns the gemini’s answer. I have then routed these function to create a flow of program,compiled it and printed it to ensure correctness, last cell is the use case of it , below is the demo run of it , as inputs are not visible in notebook .

**EXAMPLE CONVO BELOW :----**

User: hello

Assistant: Hello! How can I help you today?

User: whatsup dawg

Assistant: Hey! Just hanging out. What's up with you? Anything exciting happening?

User: 5/3-1+2

Assistant: The result is: 2.666666666666667

User: weather in new delhi

Assistant: The current weather in new delhi is mist with a temperature of 28.09°C (feels like 32.81°C). Humidity is 83%.

User: whats trending in new delhi

Assistant: Okay, here's a breakdown of what's trending in fashion in New Delhi right now, based on current observations and reports:

\*\*Overall Trends:\*\*

\* \*\*Comfort and Versatility:\*\* With people spending more time at home and also wanting to transition easily to outings, comfortable and versatile clothing is still key. Think relaxed silhouettes, breathable fabrics, and pieces that can be dressed up or down.

\* \*\*Sustainability and Conscious Consumption:\*\* Eco-friendly fabrics, upcycled clothing, and brands that prioritize ethical production are gaining traction. Consumers are more aware of the environmental impact of their fashion choices.

\* \*\*Fusion Wear:\*\* Blending traditional Indian elements with modern Western styles continues to be popular. This includes kurta dresses, crop tops with sarees, and incorporating Indian prints and embroidery into contemporary designs.

\* \*\*Return of Y2K:\*\* Low-rise jeans, crop tops, baby tees, and vibrant colors are making a comeback.

\*\*Specific Items & Styles:\*\*

\* \*\*Sarees:\*\* Sarees are a classic and always in style, but modern interpretations are popular. This includes pre-draped sarees, sarees with unconventional blouses (crop tops, shirts), and lightweight, easy-to-manage fabrics.

\* \*\*Kurti Sets:\*\* Comfortable and stylish, Kurti sets with palazzo pants, straight pants, or skirts are a go-to for many.

\* \*\*Statement Sleeves:\*\* Tops and dresses with dramatic sleeves (puffed, ruffled, bell sleeves) are a way to add flair to an outfit.

\* \*\*Bold Colors and Prints:\*\* Bright, saturated colors like fuchsia, emerald green, cobalt blue, and sunshine yellow are trending. Geometric prints, floral patterns, and abstract designs are also popular.

\* \*\*Jewelry:\*\*

\* \*\*Chunky Jewelry:\*\* Bold necklaces, statement earrings, and chunky bracelets are in.

\* \*\*Oxidized Silver:\*\* A classic Indian style, oxidized silver jewelry continues to be a favorite.

\* \*\*Minimalist Jewelry:\*\* Delicate necklaces, dainty earrings, and simple rings are also popular for everyday wear.

\* \*\*Footwear:\*\*

\* \*\*Chunky Sneakers:\*\* Comfortable and stylish, chunky sneakers are a versatile option.

\* \*\*Juttis:\*\* Traditional Indian footwear, juttis are often paired with both ethnic and Western outfits.

\* \*\*Strappy Sandals:\*\* Simple, strappy sandals are perfect for warmer weather.

\* \*\*Accessories:\*\*

\* \*\*Statement Bags:\*\* Crossbody bags, tote bags, and mini bags in bold colors or unique designs.

\* \*\*Sunglasses:\*\* Oversized sunglasses and retro-inspired frames.

\* \*\*Hair Accessories:\*\* Headbands, scarves, and clips are used to add a touch of personality to hairstyles.

\*\*Influences:\*\*

\* \*\*Bollywood:\*\* Bollywood celebrities continue to be major fashion influencers, with their outfits often setting trends.

\* \*\*Social Media:\*\* Instagram and other social media platforms play a significant role in shaping fashion trends, with influencers and bloggers showcasing the latest styles.

\* \*\*Local Designers:\*\* New Delhi has a vibrant fashion scene with many talented local designers who are creating unique and innovative pieces.

\*\*Tips for Staying on Trend:\*\*

\* \*\*Follow Indian Fashion Bloggers and Influencers:\*\* Look for those based in Delhi or who focus on Indian fashion.

\* \*\*Browse Online Retailers:\*\* Check out popular Indian e-commerce sites to see what's being promoted.

\* \*\*Visit Local Markets and Boutiques:\*\* Explore Delhi's markets (like Sarojini Nagar, Janpath, and Dilli Haat) and designer boutiques to discover unique finds.

\* \*\*Read Fashion Magazines and Websites:\*\* Stay up-to-date on the latest trends by reading Indian fashion publications.

\*\*Important Note:\*\* Fashion is subjective and personal. The best way to stay on trend is to incorporate elements that you like and that suit your individual style. Don't be afraid to experiment and create your own unique look.

Goodbye!

END OF CONVO:

Thank you cc members to allow me to work on such a great task!!!!