#### **Team Members:**

* Dagim Bantikassegn
* Manjinder Sandhu
* Tyler Dabbs
* Alex Brooks

### **1. Problem Statement**

Objective: Our goal is to organize these financial institution rules and regulations so financial institutions can go through newly created rules and regulations to know which departments to send the information to and to allow easy retrieval of these regulations when necessary. Text mining and information retrieval will allow a large step of the process to be automated.

Motivation: Automating these steps will allow less time and effort to be contributed for financial institutions employees. This allows more time and work to be used to accomplish other projects.

### **2. Data**

Dataset Overview:

* The eCFR is a website used to update laws regularly and allow anyone to be able to see new laws and regulations put into place. The eCFR dataset is based solely on Title 12 “Banking and Banks.” It is organized as a JSONL file which includes laws and regulations that banks, credit unions, and other financial institutions are required to follow. The data includes chapters, each about different subjects for financial institutions. These chapters have parts and each part is separated by “(a), (b), …”
* The preprocessing we used was first divided the data into the train and the evaluation set. This allowed us to create our golden evaluation data set later on in the model.

Annotated Data:

* The first step we took in the annotation process was for each team member to manually classify 200 texts. To do this we used ChatGPT to create 5 categories to use in the classification, and we used Ngrok so that all of us could classify our 200 texts. Using each of our 200 texts we were able to create a “gold standard” dataset.
* Using our “gold standard” dataset allows us to have strong evaluation data to determine how well our models are performing.

### **3. Model Procedure**

Model Selection:

* Earlier we discussed how we originally decided to create an NER model. We had multiple discussions about it, however, we ultimately decided that text classification would be the better fit. This is because of how diverse the dataset is. The diversity of the dataset made us feel like we would have way too many different topics. We looked through the chapter list and felt that maybe we should reconsider our options. After reevaluating the use case, we decided to switch directions toward text classification. We found that text classification would be much more accurate as text classification uses all text in each line, whereas NER uses words and phrases. Text Classification will also allow us to be able to gather better categories for the data itself.

Implementation Details:

* For Text Classification, we want to be able to create a train dataset and a test dataset. Because of our work looking at the different chapters and subchapters, we realized that we needed a way to spread the dataset out (this can be found in appendix).
* To create our categories we used ChatGPT (prompt in appendix), we asked ChatGPT for descriptions as well to allow people with less experience or knowledge in the financial industry to get a better understanding of what we are looking for.
* Using the created categories, we used Ngrok to allow us 4 to connect and classify each of the 200 texts for our evaluation or “test” set. We then created a golden evaluation set based on all our answers. This allowed us to evaluate the models we created.
* After preprocessing the jsonl file with all the data, we created a spacy model that classifies all the data with the following labels: “Capital Requirements”, “Consumer Protection'', “Risk Management”, “Reporting and Compliance”, and “Corporate Governance”. The model goes through each line/entity in the file and uses natural language processing techniques to label each regulation from the eCFR Title 12.
* We trained the model with two datasets one using 200 labels created by our team members and another using weak labels based off his labels.
* Finally, we used the model and tested it against the golden evaluation set that we all created together. This allowed us to test how well the model is at categorizing all of the texts based on the model.

### **4. Evaluation**

Quantitative Metrics:

* We used multiple metrics for all of the different categories and for the model overall. You can see the results in the appendix. Here are what each of the metrics are:
  + Accuracy: percent the model was correct.
  + Precision: Shows of all the texts the model labeled as positive how correct was the model
  + Recall: Shows of all the texts that were positive, how many were correctly labeled.
  + F1-Score: Combines both Precision and Recall to allow the calculation to include both precision and recall.

Qualitative Analysis:

We choose our 5 examples by proposals that the Board of Governors of the Federal Reserve has asked for public comment about. It would seem that when a text is very long the evaluation metrics on our model tend to top out at around .6 for accuracy, precision, recall, and f1. This could indicate that perhaps there are not enough annotated texts in our train dataset for the model to confidently classify long texts. However, shorter texts the model seems to have a much easier time with. Precision seems to consistently be the highest scoring metric meaning that our model more often than not classifies the predicted positive cases correctly. F1 score seems to score either just as high or slightly lower than both precision and recall. This means that our model overall is balanced well but could potentially have room to improve.

Results:

* Experiment 1 (we called “train200”) was created by using annotations one of our team members did themselves to allow us to train a model on these. Experiment 3 (we called “train4465”) was done from weak labels we used through the model with the train200. Both of the model’s scores can be found in the appendix. We can see that experiment 1 scored a bit better, this is probably due to being created solely by one of the groupmates instead of weak labels.

### **5. Code and Reproducibility**

Code Overview:

* You can within the GitHub Repository link all of the different python code files that are used to connect with SpaCy. With these files the user can run terminal commands to allow themselves to build the entire model and also evaluate it. The code goes through and creates a new final train to label all of the train data. It will also create “my\_trained\_model” which is the default SpaCy uses to save the dataset.

Reproducibility:

* To produce our results others need to install everything found in our “requirements.txt” file found on github. You will also need to clone the github link to your local machine. For more information see the “README” file found on the github in the appendix. We also have a project.yml file that allows us to run this code with several terminal commands.
* Terminal commands required to run are found in the appendix.

### **6. Next Steps**

Future Work:

* Some ideas to possibly add improvements to enhance the model include:
  + A validation set, this would allow us to test but also use it to validate the model before we use it to test real work.
  + Further research on other possible categories would be another great way to enhance it. We believe if our model was used by financial institutions it would be great to get their input about the five categories. This would allow us to see what new categories to add or how we should adjust the categories. This input can be gathered before production allowing us to create a more useful model.
* This model should be fine in the deployment as everything the model will work on is legally public information, because of this there is no privacy in the data itself. The biggest drawback for the model in the corporate world would be how the model works
* Recommendations:
  + One way that we could improve the model would be to possibly test on other Financial Institutions regulations. We could test on Federal Reserve data laws as well, creating an evaluation set. We can then see how the model works with Federal Reserve laws and see if we need to make any changes.
  + Another recommendation to improve the model would be to be able to take the “train200” dataset and use ngrok to get more people to create a train set. This would allow us to train data on tons of different people’s perspectives. It would be great to possibly ask people in different roles in the banking field to get all of their different perspectives.
  + We can also try improving the model using hyperparameter tuning on the spacy model by changing the n\_iter or the learn\_rate. This would be done on a trial and error basis looking for the perfect hyperparameters.

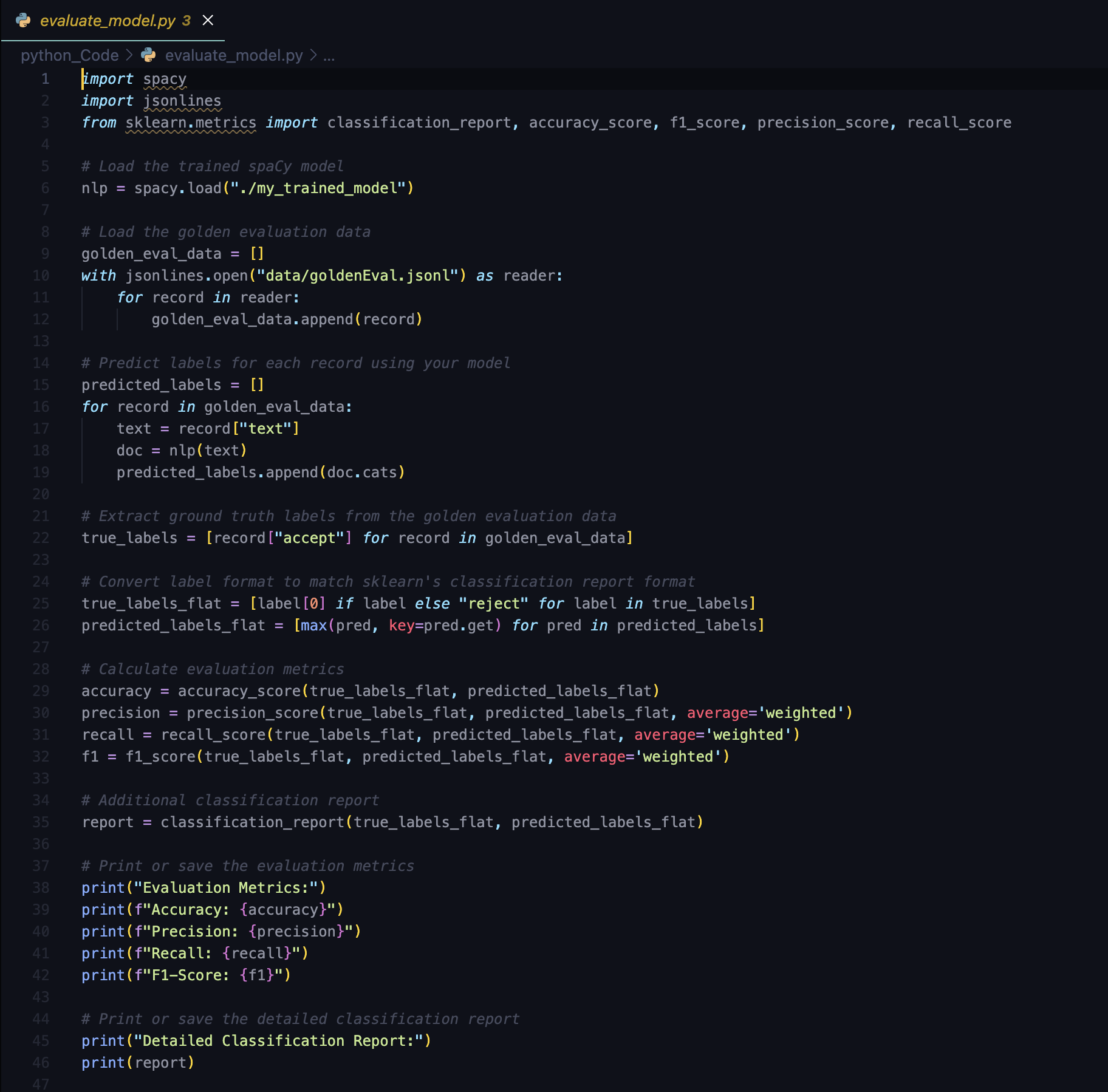
### 

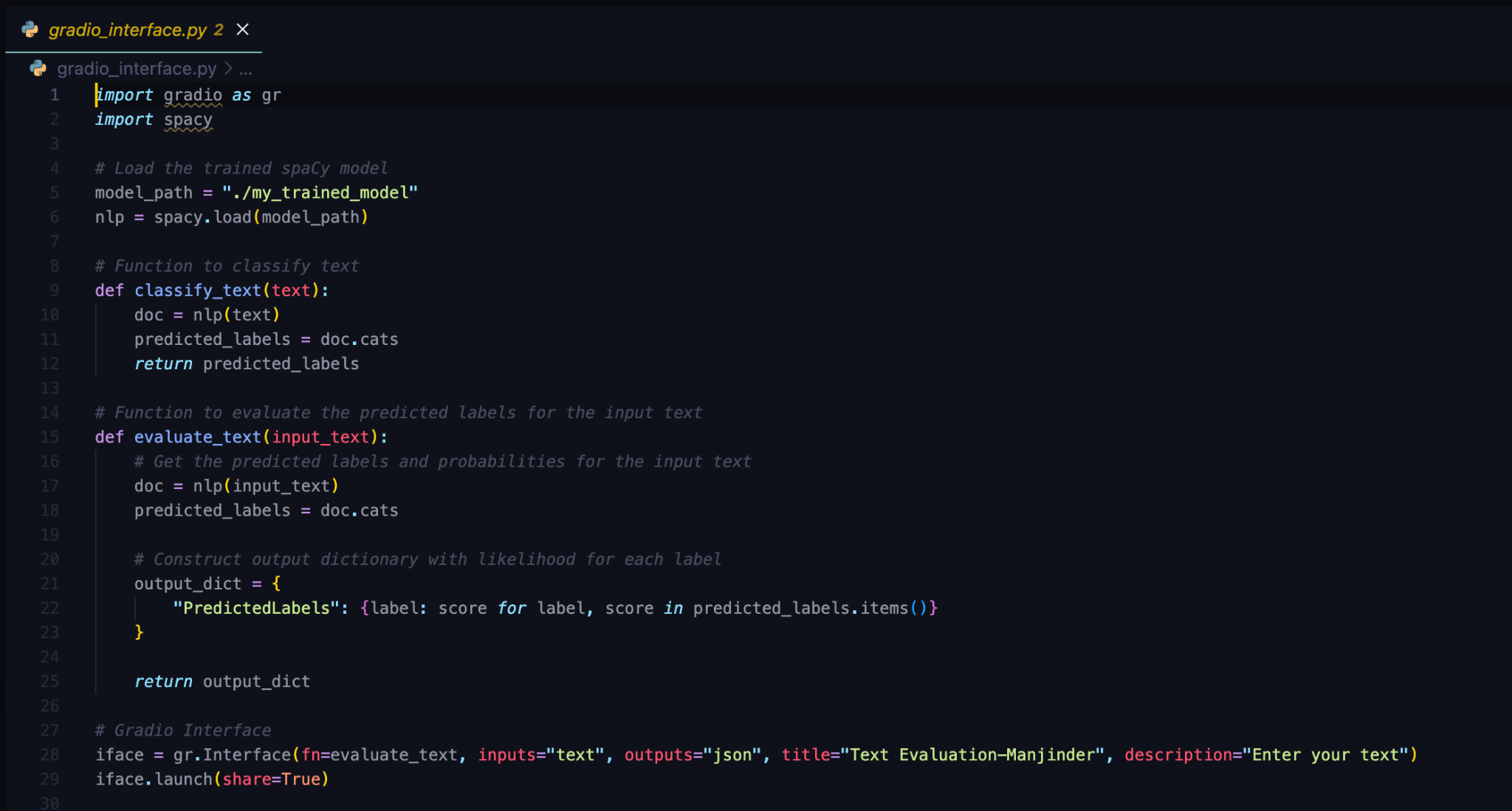
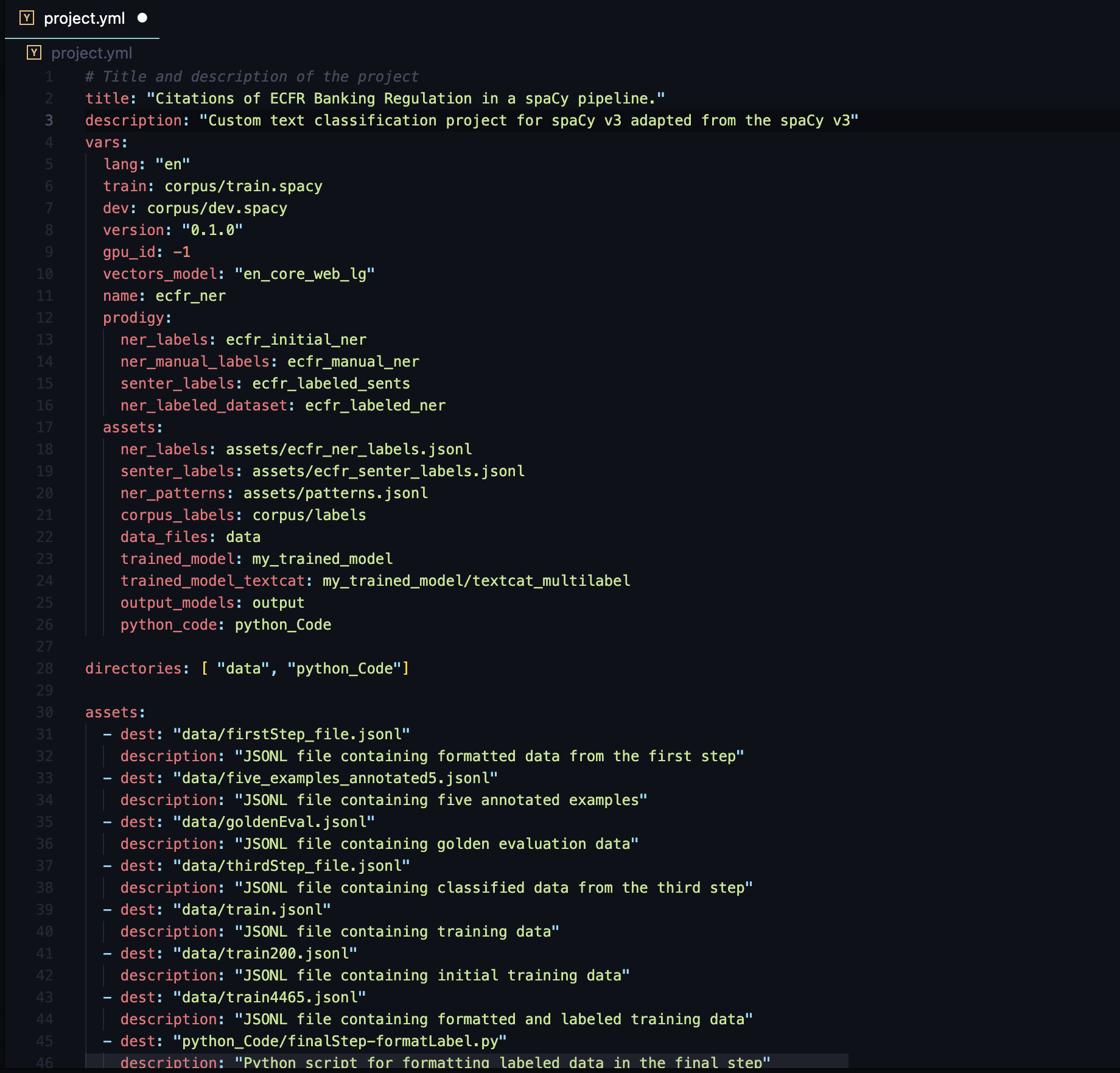
### **Appendix**

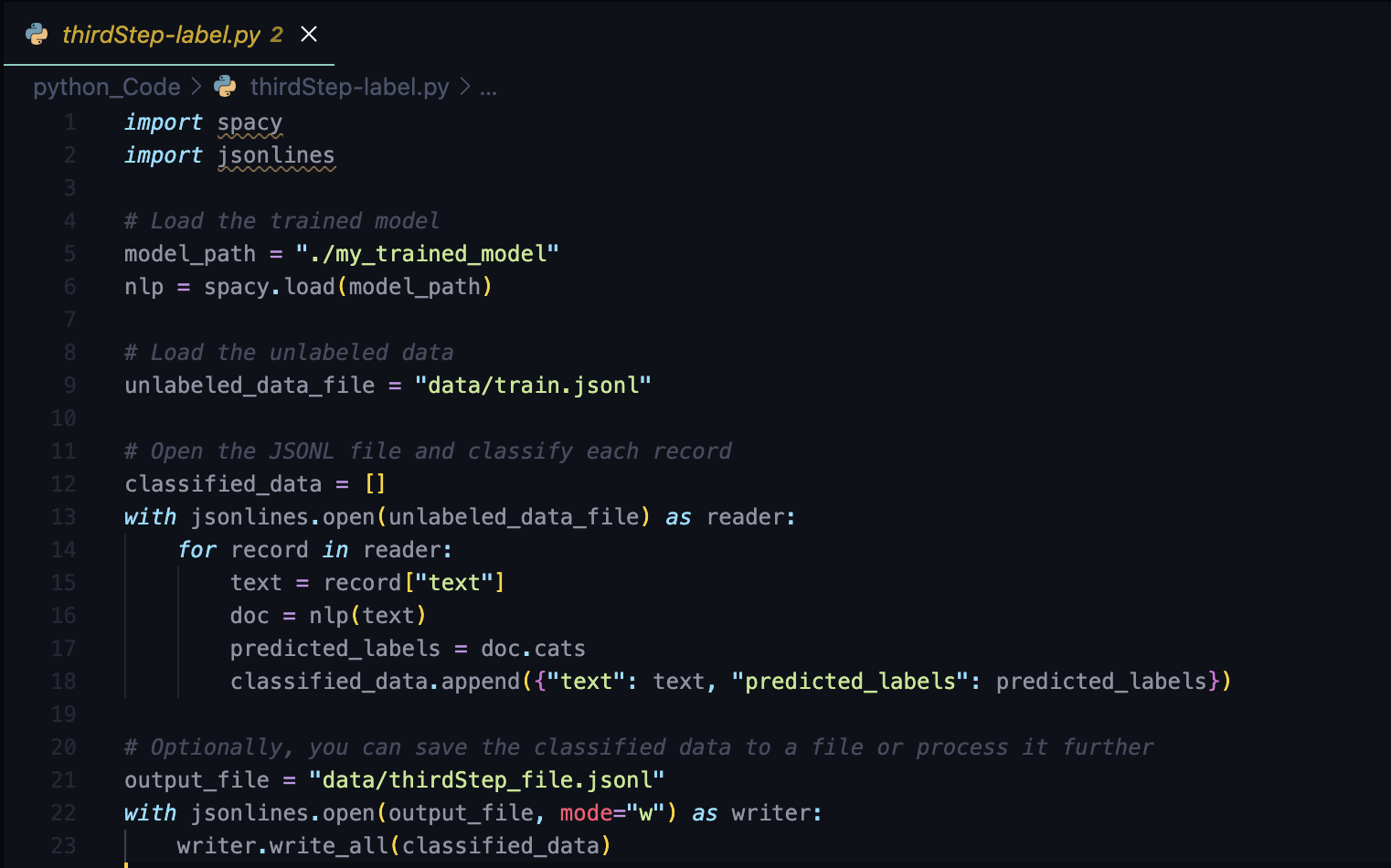
A. Code Listings

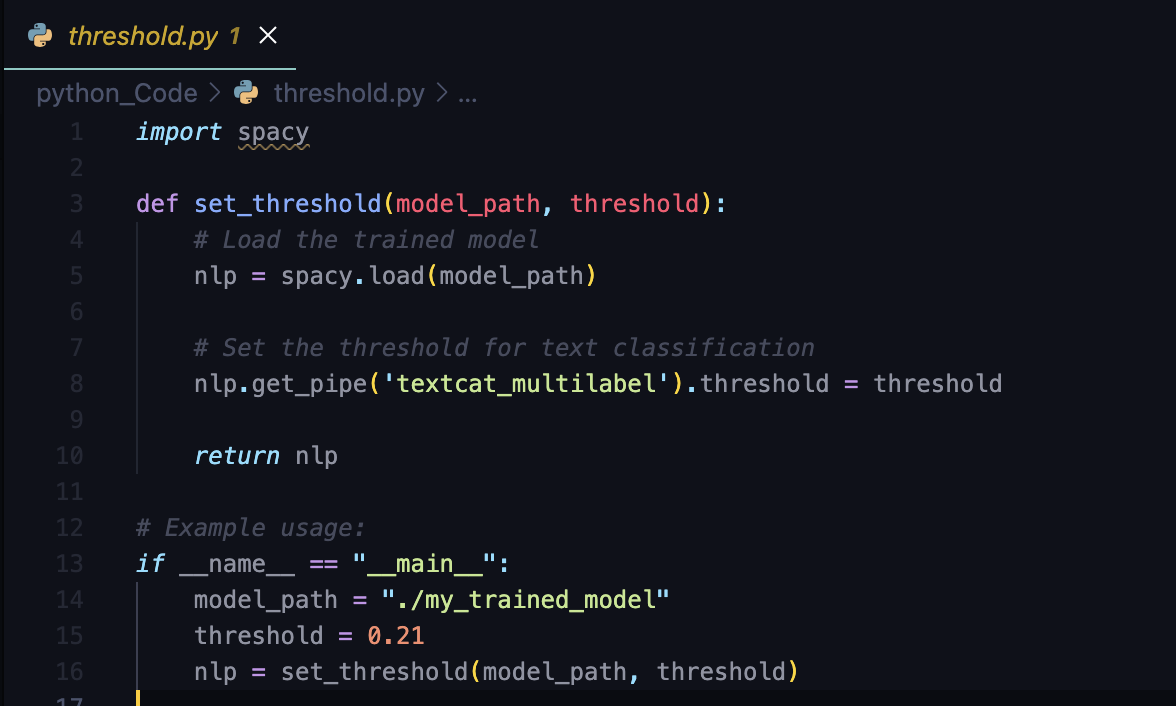
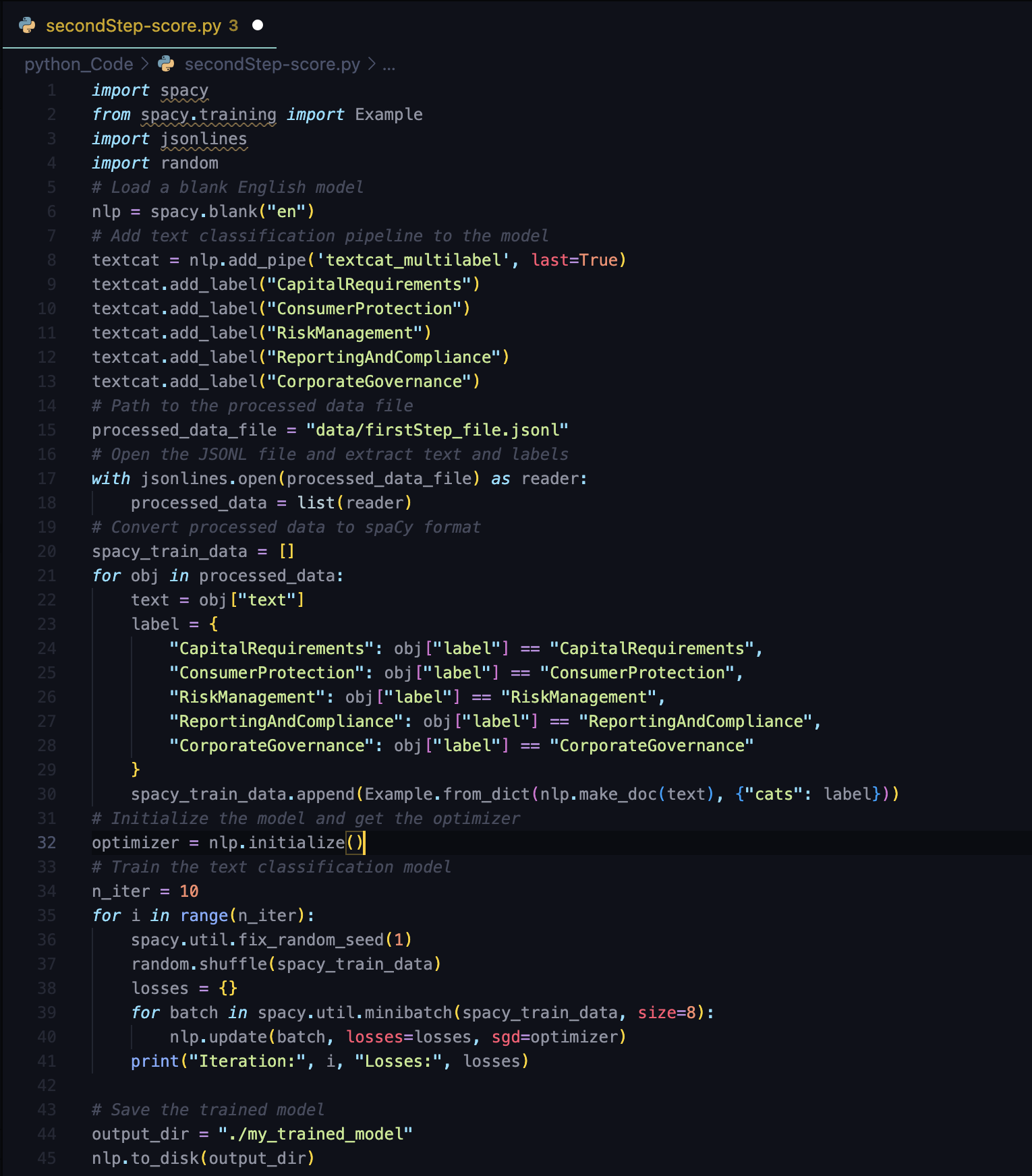
ChatGPT was also used in helping and debugging the code.

<https://github.com/ManjinderSinghSandhu/prodigy-ecfr-textcat>









In order to run this file you would have to use 2 commands in the terminal:

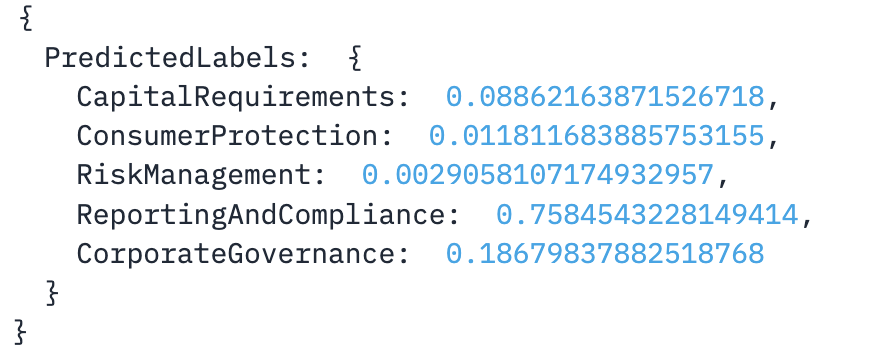
spacy project run train

spacy project run evaluate

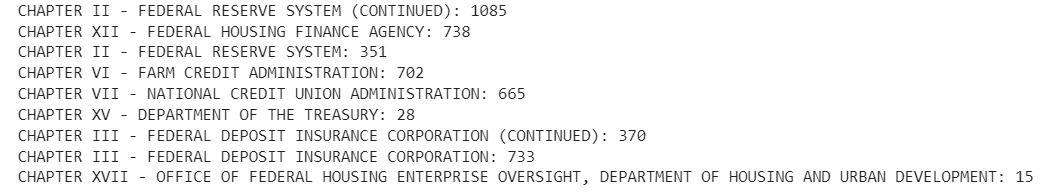
[Readme File](https://github.com/ManjinderSinghSandhu/prodigy-ecfr-textcat/blob/main/README.md)

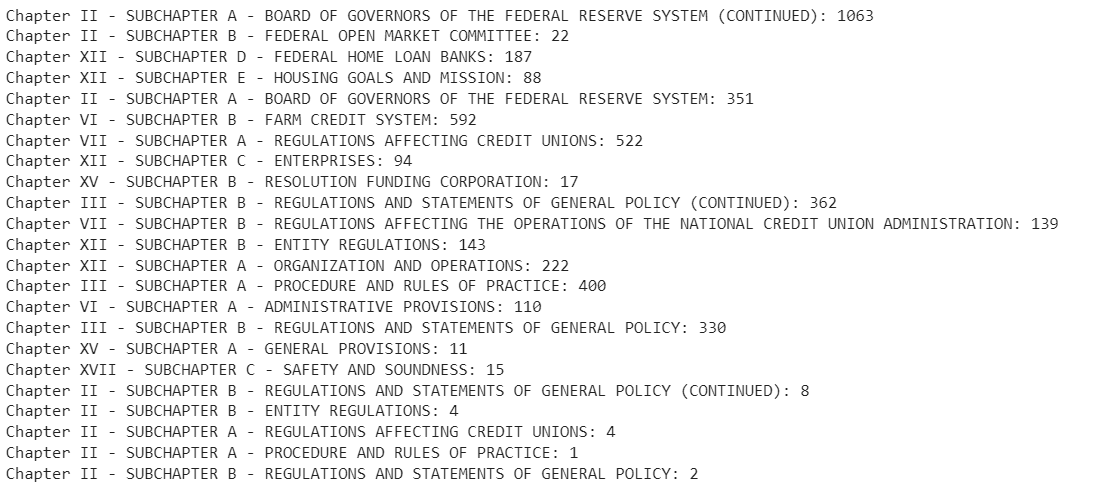
B. Inference examples

<https://huggingface.co/spaces/ManjinderUNCC/prodigy-ecfr-textcat>

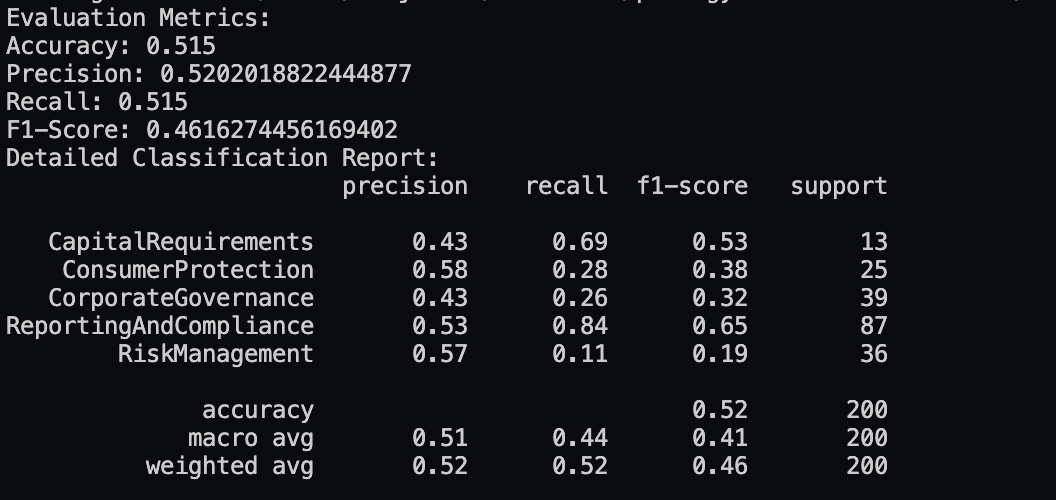
1. Text: Banks that are at risk of failing selling bonds? Absolutely not! The idea of where this money needs to come from should've been a thought that was had before these institutions took on crazy amounts of leverage and debt they couldn't pay. It's an obvious attempt at shifting the massive risk they hold onto unsuspecting investors instead of owning the bag themselves, and admitting they had no real risk management. Free money is becoming a thing of the past, it's time for these institutions to grow up and learn. Failure is always an option. Funds raised by selling off these bonds has a high chance of being similarly mismanaged by these at risk of failing institutions due to the aforementioned lack of real risk management. Actions speak louder than words, and we still live in the shadow of a great financial crisis (hmm, I wonder who could've caused that and why?) And constantly throwing the average Joe under the bus does a pretty bad job of helping maintain public confidence in the finance system. 
2. Text: The Wisconsin Bankers Association (WBA) is the largest financial trade association in Wisconsin, representing over 200 state and nationally chartered banks, savings banks,and savings and loan associations located in communities throughout the State. WBA appreciates the opportunity to comment on the interim final rule. Over the past year, the Board of Governors of the Federal Reserve System (FRB) issued several interim final rules to except certain loans that are guaranteed under the Small Business Administration's (SBA's) Paycheck Protection Program (PPP) from the requirements of the Federal Reserve Act and the corresponding provisions of Regulation O.To reflect the latest program extension by Congress, FRB issued this interim final rule to extend the Regulation O exception to PPP loans through March 31, 2022. WBA filed comment letters in support of FRB's previous interim final rules as the removal of Regulation O obstacles through the exception has helped allow Wisconsin's banks to more efficiently address the needs of their insider-owned small businesses. FRB'spast interim final rules have helped ensuree ligible businesses have timely access to liquidity to help overcome economic hurdles resulting from the effects of COVID-19 and the mitigating efforts in effect throughout Wisconsin. WBA appreciates FRB's actions to provide continued clarity that loans made by a bank to insider-owned businesses that are guaranteed under SBA's PPP remain excepted from the Federal Reserve Act and the corresponding provisions of Regulation O. Without an extension of the exception, WBA fears some auditors and examiners would treat such loans differently than PPP loans made on or before June 30 ,2020. As have been requirements of the program since inception, any PPP loan made during the extended program period must still meet certain eligibility and documentation criteria, and have the same interest rate, payment, and loan term. Additionally, all eligibility and documentation criteria and all loan terms and program requirements remain exclusively set by SBA and cannot be altered by the lender. Therefore, FRB should once again extend its exception for PPP loans; this time for PPP loans made through March 31 ,2022. WBA also appreciates FRB's efforts to have promulgated the interim final rules in such a straight-forward manner and for using plain language in its interim final rules. WBA encourages FRB to continue such efforts in future rule makings and for any other regulatory review efforts.
3. Text: How about you crooks focus on the billions being laundered by banks in plain fucking sight instead of intruding in our lives more. Disgusting.
4. Text: This proposal [R-1726], if adopted, would prove to be an invasion of privacy. In terms of digital assets, crypto exchanges are not held accountable in the same way that other financial institutions are, and have a track record of bad operational security when it comes to securely storing client information.
5. Text: Amendments to 20402(d)(2) and 204.2(e)(2) and (4) make a savings account without transfer/withdrawal limits transaction accounts. Can a depository institution avoid having a savings account be a transaction account by imposing a transfer/withdrawal restriction? Must such a restriction be absolute, or can it be suggested though the imposition of transaction fees for excess transfers/withdrawals in a stated period? The prefatory text, including the FAQ found there consistently uses the verb 'suspend.' Is 'suspend' used in the dictionary sense of 'temporarily prevent from continuing or being in force or effect'? If so, is that deliberate so as to suggest that it's expected that depository institutions will re-impose transfer/withdrawal limits at some future date (e.g., once the local economy recovers from the present pandemic)? Does the Board anticipate reinstating savings account transfer limits in the future, or believe that they will be reimposed by depository institutions as an account or contract provision? Relationship to Regulation CC A related question regarding the impact of the Reg D changes on the definition of 'account' in Regulation CC (12 CFR Part 229), which appears in the definition to exclude, except for the purposes of subpart D, any savings account described in 12 CFR 204.2(d)(2) 'even though such accounts permit third party transfers.' I note that the Official Interpretations applicable to the 229.2(a)(1) definition of 'account' in Regulation CC suggests that savings deposits are excluded because they :'may have limited third party payment powers,' and the Board believed the 'EFA Act is intended to apply only to accounts that permit UNLIMITED (emphasis added) third party transfers.' Will, then, a bank that 'suspends' its limits on savings deposit transfers and withdrawals be perforce (and perhaps unwittingly) making those savings accounts subject to Regulation CC, or does the Regulation CC 'account' definition continue to exclude savings accounts as described in 204.2(d)(2)? Thank you for your consideration of these comments and questions. 

C. Additional results





ChatGPT used for categories: <https://chat.openai.com/share/35969a1f-4fa0-41d2-8a68-08492afb33e2>



**Equations:**

Precision:

True Positives (TP) + False Positives (FP)

True Positives (TP)

​

Recall:

True Positives (TP) + False Negatives (FN)

True Positives (TP)

​

F1-Score:

2×(Precision+Recall)

( Precision×Recall)

Accuracy:

Total Number of Predictions

Number of Correct Predictions

|  | F1-Score | Accuracy | Precision |
| --- | --- | --- | --- |
| Experiment 1 (train200) | 0.491 | 0.496 | 0.573 |
| Experiment 3  (train4465) | 0.475 | 0.479 | 0.566 |

C. Terminal Commands for Reproducibility

In order to run this file you would have to setup a virtual environment and use 3 commands in the terminal:

pip install -r requirements.txt

spacy project run train

spacy project run evaluate