CSC696H Project Midterm

Sunghoon Kim

1 Introduction

Forecasting financial time series aims to anticipate predictable patterns that will bring investors advantage in trading opportunities [13]. In this project, I am going to implement a supervised machine learning model that predicts the classification of a country risk. When a situation such as a national default or social unrest occurs in a specific country, the economic damage to companies doing business with the country will be considerable. The purpose of this project is to minimize the economic damage by predicting these country risk in advance through our model. Fortunately, national defaults and social unrests do not occur suddenly, and predictive events occur. Our model analyzes the news including predictive events in the countries and uses it as dataset. In order for the results of our model to lead to the decision maker's judgment, our model must be able to explain to the users to understand the result of country risk. Research for explainability of our model will also be conducted.

1.1 OECD Country risk classification

The aims of our project is to predict Organisation for Economic Co-operation and Development (OECD) Country risk classification. OECD has the country risk classification methodology that a group of country risk experts from Export Credit Agencies(ECA) meets several times a year to update the list of country risk classifications [8]. Since the OECD's methodology uses a mixture of qualitative and quantitative models, it is difficult to predict the risk classification. This classification is used by ECA, so it becomes the major feature to evaluate importer's country risk, financial supporting export companies. If the OECD classification is downgraded, exporting companies may be provided with unfavorable conditions or may not be provided with the conditions themselves when they apply to ECA for financing necessary for export.

1.2 GDELT

When estimating country risk, I wanted to refer the news with the geopolitical crisis or natural disaster. However, in order to designate and analyze specific media companies, it was difficult to determine whether the media companies would cover all the events of all countries, and what priority to process for overlapping events between media companies. So I decided to use an open source project called GlobalData on Events, Location, and Tone (GDELT) as a dataset of our model. The GDELT Project monitors the world's broadcast, print, and web news from nearly every corner of every country in over 100 languages and identifies the people, locations, organizations, themes, sources, emotions, counts, quotes, images and events driving our global society every second of every day, creating a free open platform for computing on the entire world [6].

• *Sunghoon Kim is a graduate student at the University of Arizona. E-mail:poizzonkorea@email.arizona.edu.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

1.3 Approach of this project

Our project will derive a machine learning model to predict OECD country classification by referring to the refined news dataset provided by GDELT and economic indicators provided by the International Monetary Fund (IMF) [2] or World Bank [3]. This model is also configured with Explainable Artificial Intelligence(XAI), to provide the user with a rationale for judging our model. As trials in the related work, I plan to provide the explainability first by using post-hoc explainability models such as LIME and PDP, and to study how our model directly provides the explainability in next step, depending on the progress. In this project, it will be the first case of applying 1) XAI to country risk prediction and 2) country risk prediction for all OECD countries.

2 Related work

The related work was analyzed by dividing it into 4 areas. The contents of the OECD country risk are described in 2.1, and the machine learning model predictions for the country risk indicators are described in 2.2. Also, the project using GDELT is described in 2.3, and the XAI work for credit risk is described in 2.4.

2.1 OECD Country Risk

As shown in the figure below, the OECD publishes country risk classification reports for 201 countries at least twice a year. It has been published 93 times from 1999 to March 2022 and the most recent publish was made on March 11, 2022.

Country Risk Classifications
of the Participants to the Arrangement on Officially Supported Export Credits
Valid as of: **11 March 2022**

			Classification		
nb	Country Code ISO Alpha 3	Country Name [1]	Previous	Current Prevailing	Notes
1	AFG	Afghanistan	7	7	
2	ALB	Albania	5	5	
3	DZA	Algeria	5	5	
4	AND	Andorra	-	-	(9)
5	AGO	Angola	6	6	
6	ATG	Antigua and Barbuda	7	7	(8)
7	ARG	Argentina	7	7	
8	ARM	Armenia	6	6	
9	ABW	Aruba	6	6	
10	AUS	Australia	-	-	(6)
11	AUT	Austria	-	-	(6) (7)
12	AZE	Azerbaijan	4	4	
13	BHS	Bahamas	4	4	
14	BHR	Bahrain	6	6	
15	BGD	Bangladesh	5	5	
16	BRB	Barbados	-	-	(5)
17	BLR	Belarus	6	**7**	
18	BEL	Belgium	-	-	(6) (7)
19	BLZ	Belize	-	-	(5)

Fig. 1. Country Risk Classifications

Since both the GDELT dataset and the OECD country risk use the ISO country code, it is possible to link the two datasets. The national ratings are distributed from 1st to 7th grade, where 1st grade is a high-reliability country and 7th grade is a low-reliability country. In the most recent publish, two countries had their credit ratings downgraded, with Russia and Belarus both downgrading to 7th. Some countries are not evaluated for reasons such as "High Income OECD or Euro Area Country not reviewed or classified." or "Currently not reviewed or classified.", so I will exclude them from the forecast in this project as well.

2.2 ML Model for country risk indicators

"Multivariate CDS risk premium prediction with SOTA RNNs on MI[N]T countries" [5] predicts CDS risk premiums of Mexico, Indonesia and Turkey by applying state-of-the-art forecasters in deep learning recurrent neural networks architectures. A credit default swap (CDS) is a financial swap agreement that the seller of the CDS will compensate the buyer in the event of a debt default (by the debtor) or other credit event [10]. In short, an increase in the CDS premium indicates an increase in default risk, and a decrease in the CDS premium indicates a decrease in default risk. The architectures used in this article, RNN (ELMAN), NARX, GRU, and LSTM, will be applied as candidates to be the best learning architecture in this project as well. In order to check the learning performance, like this article, I'm going to use MSE, MAE, and R^2(=Root Mean Square Error) as a criterion of the performance. Also, in this article, each country has different optimized machine learning architectures, and as in our project, it is expected that the dominant architecture will be easily distinguished as the number of country set increases.

"Are CDS spreads predictable during the Covid-19 pandemic? Forecasting based on SVM, GMDH, LSTM and Markov switching autoregression" [13] investigates the forecasting performance for CDS spreads by Support Vector Machines (SVM), Group Method of Data Handling (GMDH), Long Short-Term Memory (LSTM) and Markov switching autoregression (MSA) for daily CDS spreads of the 513 leading US companies, in the period 2009–2020. Although this article targeted companies for credit ratings instead of countries, it was found that even under special circumstances such as COVID-19, the performance of the rating prediction model was not significantly affected. This implies that if the resulting model of our project is implemented well, it can be applied universally.

"Examination of Country Risk Determinants Using Artificial Neural Networks: The Case of Turkey" [11] constructs a functional model of forecasting country risk changes in Turkey with the help of artificial neural networks. This article predicts the OECD country classification of Turkey, and the correct classification rate has increased up to 98.5% for the terms with improved conditions. This made it possible to realize that the predictive model for the OECD country risk classification is feasible. In addition, the influence of each economic indicator on the model is presented as shown in the table below, giving hints in the selection of economic indicators in our project.

INCREASE IN COUNTRY RISK	DECREASE IN COUNTRY RISK
Wholesale Price Index (WPI), Market Interest Rate, Open Market Operations (OMO)	GDP, Index of Employment, Consumer Confidence Index, Real FX Rate Index, External Debt, Domestic Assets, Issue Volume (Money Emission) Reserves, Wage Index

* Risk effects are determined with respect to ascending trend in the corresponding variables.

Fig. 2. Summary on Risk Effects of Model Variables

2.3 How to handle GDELT

The GDELT Project [6] is a realtime network diagram and database of global human society for open research which monitors the world's events. GDELT dataset has 2 versions. GDELT 2.0 has 3 more fields and updates every 15 minutes, while GDELT 1.0 updates every single day, but it only manages events from February, 2015, so it doesn't fit the time-series in our project. Therefore, our project use GDELT 1.0 datasets.

GDELT 1.0 has 58 fields. In our project, I use the following four fields from a record: SQLDATE, EventBaseCode, GoldsteinScale, Actor_CountryCode. SQLDATE is the date the event took place in YYYYMMDD format. EventBaseCode denotes level two leaf root node. For example, code 1452 (engaging in violent protest for policy change) has a EventBaseCode of 145 (Protest violently, riot, not specified below). Examples of EventBaseCode and its description is written in the table below. GoldsteinScale is a numeric score from -10 to +10, capturing the theoretical potential impact that type of event will have on the stability of a country. Actor_CountryCode is the country of the actors, which is a 3-character ISO country code for the location.

Table 1. The EventBaseCode and descriptions

EventBaseCode	Description
101	Demand material cooperation
111	Criticize or denounce
121	Reject material cooperation
131	Threaten non-force
141	Demonstrate or rally
151	Increase police alert status
161	Reduce or break diplomatic relations

"Predicting Social Unrest Events with Hidden Markov Models Using GDELT" [9] builds a Hidden Markov Models (HMMs) based framework to predict indicators associated with country instability using autocoded events dataset GDELT. Through this article, our project got hints on which fields to extract from GDELT. "Estimating countries' peace index through the lens of the world news as monitored by GDELT" [12] predicts countries' peace index by Elastic Net, Decision Tree, and Random Forest, and the process of their dynamic training. This article describes in detail the difference between Tone and Goldstein among GDELTs. Tone is a measure of emotional impact, and the Goldstein scale is a measure of potential impact. For example, an event such as "Decline comment" indicates a lower value in Tone, and an event such as "Threaten with military force" indicates a lower value in the Goldstein scale. In our project, I decided to use the Goldstein scale, judging that the potential impact will have a more significant impact on the country credit rating.

2.4 XAI in credit risk

"Towards Explainable Deep Learning for Credit Lending: A Case Study" [7] explored the process of explaining credit lending decisions made by a neural network using three different attribution methods: LIME, DeepLIFT, and Integrated Gradients. "ENABLING MACHINE LEARNING ALGORITHMS FOR CREDIT SCORING" [1] showed how to take credit scoring analytics in to the next level, namely they present comparison of various predictive models (logistic regression, logistic regression with weight of evidence transformations and modern artificial intelligence algorithms) and show that advanced tree based models give best results in prediction of client default. Both documents apply the post-hoc explainability model to credit risk machine learning model, making a business case viable for now "not-so-black-box models". Our project will also consider applying such a surrogate model as the primary goal, and expansion of the direct explainability model will also be considered. In particular, in the case of Experiment 1 of [7], it seems that the idea can be applied to our project because an evaluation method that excludes humans' judgement is implemented in evaluating XAI.

3 PLAN

The project will be carried out in three steps. The first step is to focus on setting up the development environment, applying the GDELT dataset to our project, and implementing a vanilla prediction model using LSTM. The second step is to improve the performance of the predictive model and analyze the performance of our model. The final step is to conduct an experiment on XAI.

(Step 1) Instead of the cloud system, this project use the local development environment. I tried to access GDELT using BigQuery provided by Google, but it took longer than expected to load the initial data, so I gave up. Instead, data from 1998 to March 11, 2022 from GDELT were downloaded locally and the pivot table for each year was created in the form of a csv file. Using this file as a dataset and Keras library in Python, I plan to implement a basic supervised machine learning model using LSTM by the end of March.

(Step 2) Similar to many machine learning projects, this project find optimized performance using randomized hyperparameter tuning. In addition to RNN, this project will derive an optimal learning model by experimenting with predictive models using RNN, GRU, and MSA. Also, I will compare the prediction performance when economic indicators such as exchange rates, interest rates, and stock indices are combined to the GDELT dataset.

(Step 3) First, this project test explainability using post-hoc models, LIME and PDP. The model results will be shared with ECA staff to receive feedback. Also, the applicability of some of the direct explainable models presented in class will be reviewed.

Table 2. Project Milestones

D-day	Description
3/31	Prepare for GDELT dataset
	Implement Valilla model using LSTM
4/15	Refine the model through hyper-parameter tuning
	and selection of appropriate learning architectures
	Experiment with combinations of GDELT + economic indicators
4/30	Apply post-hoc explainiability model
	Evaluate XAI by users' feedback and computation model
	(Option) Apply direct explainiability model

4 Results

Vanilla model implementation will start from the example source code provided by Keras's "Timeseries forecasting for weather prediction" [4]. This is because it is the source code for prediction using timeseries data that is most similar to the objective of our project. Randomized hyper-parameter tuning will refer to the project results of CSC580, which is my previous machine learning project. Also, refer to "Multivariate CDS risk premium prediction with SOTA RNNs on MI[N]T countries" [5] to find the optimized learning model among RNN, GRU, LSTM, and MSA, and refer to "Enabling Machine Learning Algorithms for Credit Scoring" [1] to check the performance through the combination of economic indicators. As with [1], if a guide message in the form of the table below appears, it is expected that the project will be successful.

Sample guide messages for personal credit risk

- please, pay on time as any delay in instalment payment 30 days will significantly decrease of any banking financing in the future,
- please, do not exceed the limit of your revolving facilities,
- please, reduce the usage of the revolving facilities, as the high usage negatively impacts your scoring,
- please, develop your credit history by applying for a card or a loan and servicing it regularly.

References

[1] P. Biecek, M. Chlebus, J. Gajda, A. Gosiewska, A. Kozak, D. Ogonowski, J. Sztachelski, and P. Wojewnik. Enabling machine learning algorithms for credit scoring–explainable artificial intelligence (xai) methods for clear understanding complex predictive models. *arXiv preprint arXiv:2104.06735*, 2021.

[2] I. M. Fund. Imf data.

[3] W. B. Groupd. World bank data.

[4] Keras. Timeseries forecasting for weather prediction.

[5] Y. Kutuk and L. Barokas. Multivariate cds risk premium prediction with sota rnns on mi [n] t countries. *Finance Research Letters*, p. 102198, 2021.

[6] K. Leetaru and P. A. Schrodt. Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, vol. 2, pp. 1–49. Citeseer, 2013.

[7] C. Modarres, M. Ibrahim, M. Louie, and J. Paisley. Towards explainable deep learning for credit lending: A case study. *arXiv preprint arXiv:1811.06471*, 2018.

[8] OECD. Oecd country risk classification.

[9] F. Qiao, P. Li, X. Zhang, Z. Ding, J. Cheng, and H. Wang. Predicting social unrest events with hidden markov models using gdelt. *Discrete Dynamics in Nature and Society*, 2017, 2017.

[10] the Wharton School of the University of Pennsylvania. Cdos are back: Will they lead to another financial crisis?

[11] M. S. Topak and E. Muzir. Examination of country risk determinants using artificial neural networks: The case of turkey. *International Research Journal of Finance and Economics*, 75(120), 2011.

[12] V. Voukelatou, L. Pappalardo, I. Miliou, L. Gabrielli, and F. Giannotti. Estimating countries' peace index through the lens of the world news as monitored by gdelt. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pp. 216–225. IEEE, 2020.

[13] D. B. Vukovic, K. Romanyuk, S. Ivashchenko, and E. M. Grigorieva. Are cds spreads predictable during the covid-19 pandemic? forecasting based on svm, gmdh, lstm and markov switching autoregression. *Expert systems with applications*, 194:116553, 2022.