International Review of Financial Analysis 82 (2022) 102174

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Economic policy uncertainty and bankruptcy filings

Elena Fedorova[a,b], Svetlana Ledyaeva[c,*], Pavel Drogovoz[d], Alexandr Nevredinov[d]

[a] Financial University, Department of Corporate Finance and Corporate Governance, Moscow, Russian Federation
[b] National Research University Higher School of Economics (HSE), Department of Finance, Moscow, Russian Federation
[c] Aalto University School of Business, Department of Economics, Helsinki, Finland
[d] Bauman Moscow State Technical University, Department of Engineering Business and Management, Moscow, Russian Federation

ARTICLE INFO

Keywords:
economic policy uncertainty, bankruptcy, firm, machine learning, stacking, bagging

ABSTRACT

Applying machine learning techniques to predict bankruptcy in the sample of French, Italian, Russian and Spanish firms, the study demonstrates that the inclusion of economic policy uncertainty (EPU) indicator into bankruptcy prediction models notably increases their accuracy. This effect is more pronounced when we use novel Twitter-based version of EPU index instead of original news-based index. We further compare the prediction accuracy of machine learning techniques and conclude that stacking ensemble method outperforms (though marginally) machine learning methods, which are more commonly used for bankruptcy prediction, such as single classifiers and bagging.

1. Introduction

From the second half of 20th century, business failure has emerged as an extensively researched area (Kumar & Ravi, 2007). The reason for it is a common interest of many economic and financial actors. The prediction of bankruptcy is one of the most important business decision-making problems facing auditors, consultants, management, banks, and government policy makers (O'Leary, 1998). Although the classical studies dedicated to bankruptcy issues used only accounting and market data (Altman, 1968; Beaver, 1966; Fitzpatrick, 1932), over time it has become clear that any model which contain only financial statement information will not predict accurately failure or nonfailure of a firm (Zavgren, 1985). Therefore, researchers started to introduce new types of indicators in bankruptcy models. Generally, all bankruptcy-related factors could be grouped into two main categories: internal and external. In turn, internal are divided into accounting and corporate governance factors, while external – into macroeconomic and market factors.

Numerous studies confirm the dependency between macroeconomic variables (such as GDP, GDP growth, interest rate, volatility of foreign exchange rate) and bankruptcy rates (see, among others, Levy & Bar-Niv, 1987; Platt, Platt, & Pedersen, 1994; Hol, 2007; Dewaelheyns & Van Hulle, 2008; Bhattacharjee, Higson, Holly, & Kattuman, 2009; Sarikov & Kuprianov, 2020). We add to this literature by suggesting that economic policy uncertainty (EPU) can significantly affect the probability of firms` bankruptcy. Though causal relationship between economic policy uncertainty and bankruptcy has been already confirmed in Stolbov and Shchepeleva (2020), our study shows that the inclusion of EPU indicator as an external non-financial macroeconomic factor into bankruptcy prediction model notably increases its accuracy. In addition, we utilize a novel modification of commonly used EPU index proposed by Baker, Bloom, and Davis (2016). This index reflects the changes in policy-related economic uncertainty. The original index has been calculated based on newspaper coverage frequency. One of the main disadvantages of this index is a modest dataset for some countries. If the dataset for the United States contains 10 newspapers, the datasets for other countries may use only one (China, Russia) or two (France, Germany, Italy, Japan, Spain, United Kingdom) newspapers. To overcome this disadvantage, Renault, Baker, Bloom, and Davis (2021) have calculated Twitter-based EPU index which was focused on only English-speaking part of data. In this study we go further and compute Twitter-based EPU index using native-speaking part of data. This enables us to deliver more reliable comparison of performance of original and Twitter-based indices in predicting bankruptcy. This has partially determined our choice of countries for the analysis (France, Italy, Russia, and Spain) as they should have a core strong language that is not English.

There has been ample literature regarding models for predicting bankruptcy. Seminal academic research has evaluated bankruptcy using traditional statistics techniques (such as discriminant analysis and

* Corresponding author.
E-mail addresses: svetlana.ledyaeva@aalto.fi (S. Ledyaeva), drogovoz@bmstu.ru (P. Drogovoz).

https://doi.org/10.1016/j.irfa.2022.102174
Received 8 November 2021; Received in revised form 9 March 2022; Accepted 19 April 2022
Available online 22 April 2022

logistic regression) and artificial intelligence models (for example, artificial neural networks). Since 1990`s, machine learning techniques have been extensively applied to predict firms` bankruptcy (Qu, Quan, Lei, & Shi, 2019). In a recent study, Barboza, Kimura, and Altman (2017) conclude that machine learning models show, on average, approximately 10% more accuracy in relation to traditional models.

Machine learning methods applied to bankruptcy prediction include single and ensemble classifier models. Ensemble classifier models can be divided into three types: bagging, boosting, and stacking. In prior research, bankruptcy prediction models have been most often estimated using single classifiers or/and boosting or/and bagging ensembles (see Barboza et al., 2017) while the application of a stacking ensemble (an otherwise widely used machine learning technique) to bankruptcy prediction has not been fully explored (Liang, Tsai, Lu, & Chang, 2020). There is no consensus whether stacking ensemble performs better or worse than other ensemble methods in predicting bankruptcy. For example, Kim (2018) and Liang, Tsai, Dai, and Eberle (2018) show that stacking ensembles outperform bagging and boosting ensembles, while Pisula (2020) comes to an opposite conclusion. In this study, we apply stacking ensemble technique for bankruptcy prediction and compare its performance with single classifier and bagging ensemble models.

Our contribution to the literature is two-fold. First, we show that the inclusion of economic policy uncertainty indicator (particularly its novel Twitter-based version as compared to original news-based index) into bankruptcy prediction models notably increases their accuracy. This way the study also contributes to emerging literature on the role of social media in firm-level bankruptcy/financial distress prediction (Dunham & Garcia, 2021; Jabeur, Stef, & Carmona, 2022; Putra, Joshi, Redi, & Bozzon, 2020). Second, we acknowledge that though in our estimations, on average, stacking ensemble models show higher accuracy in bankruptcy prediction in relation to other machine learning models tested in this study (single classifiers and bagging ensemble), this difference is not remarkable.

The rest of this paper is organized as follows. Section 2 connects the literatures on bankruptcy prediction and economic policy uncertainty and outlines theoretical framework of the study. Section 3 describes the data and variables while fourth section introduces method and research roadmap. Section 5 presents the results and Section 6 concludes the paper.

2. Background

2.1. Internal and external factors of bankruptcy

The literature on bankruptcy prediction dates back to the 1930's beginning with the studies concerning the use of ratio analysis to predict business failure. Until the mid-1960`s research was focused on univariate (single factor/ratio) analysis (Bellovary, Giacomino, & Akers, 2007). Altman (1968) has been the first multivariate study, which remains very popular at present times. Up to date, there is a great variety in bankruptcy prediction models regarding how many and which factors are considered in them. Initially, bankruptcy factors reflected only internal (i.e. company-level) business and financial activities. Bellovary et al. (2007) list 42 most common internal factors in bankruptcy prediction models. The top ten factors include *Net income/Total assets*, *current ration*, *Working capital/Total assets*, *Retained earnings/Total assets*, *Earning before interest and taxes/Total assets*, *Sales/Total assets*, *Quick ratio*, *Total debt/Total assets*, *Current assets/Total assets*, and *Net income/ Net worth*.

However, over time researchers started to question the accuracy of a bankruptcy prediction model which consider only internal financial information (Zavgren, 1985). Therefore, scholars began to introduce new types of factors in bankruptcy models, particularly external or macroeconomic factors. Historically, bankruptcy filings have closely followed general economic conditions as businesses and households seek relief from macroeconomic shock (Wang, Yang, Iverson, & Kluender, 2020). Hence, accounting for changes in macroeconomic conditions is important in assessing the probability of the bankruptcy filing option. Previous research on macroeconomic/external factors of bankruptcy filings (including bankruptcy prediction models) has focused on examining the role of general macroeconomic indicators such as GDP and its growth (for example, Dewaelheyns & Van Hulle, 2008; Hol, 2007; Levy & Bar-Niv, 1987; Santoro & Gaffeo, 2009), interest rate (for example, Ninh, Do Thanh, & Hong, 2018; Platt et al., 1994), inflation (for example, Levy & Bar-Niv, 1987; Santoro & Gaffeo, 2009) and volatility of foreign exchange rate (Nam, Kim, Park, & Lee, 2008). However, to the best of our knowledge, no prior research has studied economic policy uncertainty as a macroeconomic factor of bankruptcy filings (except Stolbov and Shchepeleva (2020) who utilize Granger causality test to study causal relationship between economic policy uncertainty and bankruptcy). In this paper we test whether the inclusion of economic policy uncertainty indicator into bankruptcy prediction models increases their accuracy.

2.2. Literature review on the effects of economic policy uncertainty

Baker et al. (2016) EPU index has already proven its efficiency in explaining various economic, financial, and business indicators, both economy-wide (macro-) and firm-level (micro-). Studies on macroeconomic effects include, for example, Antonakakis, Chatziantoniou, and Filis (2014) who conclude that oil price shocks negatively respond to EPU shocks. Istrefi and Piloiu (2014) show that inflation expectations are sensitive to policy-related uncertainty shock. Caggiano, Castelnuovo, and Figueres (2017) study the effects of an unanticipated increase in economic policy uncertainty on unemployment in recessions and expansions and find that the response of unemployment to be statistically and economically larger in recessions. Abid (2020) find that exchange rates movements in emerging markets are driven to a great extent by economic policy uncertainty. Cepni, Guney, and Swanson (2020) demonstrate that the inclusion of economic policy uncertainty factor leads to superior prediction of GDP growth.

Macro financial effects of EPU have been examined in numerous studies as well. Brogaard and Detzel (2015) find that Baker et al. (2016) EPU index positively forecasts log excess market returns. Bordo, Duca, and Koch (2016) and Nguyen, Le, and Su (2020) show that higher level of EPU has negative impact on bank credit growth. Ashraf and Shen (2019) demonstrate that increase in economic policy uncertainty is associated with higher average interest rates on bank gross loans. Phan, Iyke, Sharma, and Affandi (2021) find that EPU has a negative impact on financial stability though the final effect depends on the financial system characteristics. Liu and Zhang (2015) conclude that higher EPU leads to significant increases in stock market volatility and that incorporating EPU as an additional predictive variable into the existing volatility prediction models significantly improves forecasting ability of these models.

There is also ample research on firm-level effects of economic policy uncertainty. Demir and Ersan (2017) and Phan, Nguyen, Nguyen, and Hegde (2019) find that firms prefer to hold more cash when economic policy uncertainty increases. Zhang, Han, Pan, and Huang (2015) show that, on average, Chinese firms` leverage ratios decrease when economic policy uncertainty increases. Iqbal, Gan, and Nadeem (2020) demonstrate robust evidence that the effect of EPU on firm performance is negative. Li (2020) concludes that economic policy uncertainty increases frequency and volume of insider trades.

It is important to point to the direct relation of the outlined previous literature to this study research question. In the mentioned studies EPU was proven to affect not a few macro- indicators (GDP growth, inflation, interest rates, exchange rate), which have already been used as factors of bankruptcy prediction in prior research (see, for example, Levy & Bar-Niv, 1987; Platt et al., 1994; Hol, 2007; Dewaelheyns & Van Hulle, 2008; Bhattacharjee et al., 2009; Sarikov & Kuprianov, 2020). Firm-level indicators proven to be affected by EPU have been also often considered in bankruptcy prediction models (see, for example, Laitinen and Laitinen (1998) for cash holding and Son, Hyun, Phan, and Hwang (2019) for leverage and profitability ratios). This further reinforces applicability of the inclusion of economic policy uncertainty measure into bankruptcy prediction models.

Finally, a significant direction of the prior research explores how EPU affects firm-level investment (see, for example, Kang, Lee, & Ratti, 2014; Gulen & Ion, 2016; Drobetz, El Ghoul, Guedhami, & Janzen, 2018; Liu & Zhang, 2020; Chen, Lee, & Zeng, 2019; Hou, Tang, Wang, & Xiong, 2021). The common conclusion is that economic policy uncertainty negatively affects firm-level investment. This inference is an important component of our theoretical framework which we describe below.

2.3. Theoretical framework

The theoretical research on the impact of uncertainty on firms` decisions has a long history. In his seminal paper, Bernanke (1983) builds a theoretical model to show that high uncertainty gives firms an incentive to delay investment and hiring when investment projects are costly to undo or workers are costly to hire and fire.[1] Baker et al. (2016) point out that there are multiple reasons for depressive effects of uncertainty including precautionary spending cutbacks by households, upward pressure on the cost of finance (for example, Gilchrist, Sim, & Zakrajšek, 2014; Pástor & Veronesi, 2013), managerial risk aversion (for example, Panousi & Papanikolaou, 2012), and interactions between nominal rigidities and search frictions (Basu & Bundick, 2012; Leduc & Liu, 2016).

There is also ample theoretical literature that has focused explicitly on policy (monetary, fiscal, regulatory) uncertainty and particularly its detrimental effects for firms and economy in general (Friedman, 1968; Hassett & Metcalf, 1999; Higgs, 1997; Rodrik, 1991). A common view of this research is that policy uncertainty discourages firm-level investment. To date, numerous empirical studies confirmed this theory (see, for example, Kang et al., 2014, Gulen & Ion, 2016, Drobetz et al., 2018, Liu & Zhang, 2020, Chen et al., 2019, and Hou et al., 2021).

At another point, Lyandres and Zhdanov (2013) build a simple theoretical model that shows that shareholders of a firm with valuable investment opportunities would be able to wait longer before defaulting on their contractual debt obligations compared to shareholders of an otherwise identical firm without such opportunities. They further empirically demonstrate that measures of firms` investment opportunities (firm-level, such as market-to-book ratio) are negatively related to the likelihood of bankruptcy in the data. Finally, they show that the inclusion of these measures into bankruptcy prediction models significantly improves their forecasting ability.

Abovementioned deduction allows us to propose that bankruptcy prediction models should benefit from the inclusion of economic policy uncertainty indicator as one of the core factors of firm-level investment, which in turn is an important factor of bankruptcy.

3. Data and variables

3.1. Economic policy uncertainty index

In this study we utilize Economic Policy Uncertainty index proposed by Baker et al. (2016) as an external non-financial macroeconomic factor of bankruptcy. This index reflects the changes in policy-related economic uncertainty and initially has been calculated based on newspaper coverage frequency. The index is constructed using the following procedure. First, the authors collect the data from most respected newspapers in a country and calculate the number of articles where some economic policy uncertainty is mentioned. To select economic policy uncertainty articles, special banks of words have been applied. To be considered as an uncertainty article it should contain terms from all three banks of words dedicated to: economy, policy, and uncertainty. Next, they find the share of such articles for each month and normalize obtained numbers to a mean of 100 for the entire time-series period.

Although, EPU index (which can be found on the internet resource: https://www.policyuncertainty.com/) is a promising macroeconomic indicator, it still suffers from some drawbacks. The most important is a modest dataset for some countries. If the dataset for the United States contains 10 newspapers, the datasets for other countries may use only one (China, Russia) or two (France, Germany, Italy, Japan, Spain, United Kingdom) newspapers. Hence, we decided to recalculate EPU index using a larger dataset.

Nowadays, the increasing role of online social media (Twitter, Facebook, Instagram) is indisputable. They gradually replace the traditional media at least in developed countries. Traditional and online media differ in many areas. Most particularly, the content of news shared by social networks is much more diversified because of lack of restrictions which can exist in traditional media. Moreover, it is more focused on entity-oriented topics that have low coverage in traditional news media (Zhao et al., 2011). Social network news tend to be more controversial and contain emotional features (García-Perdomo, Salaverría, Kilgo, & Harlow, 2018). That is why they may have higher impact on people's feeling and perceptions. The key audience of traditional and social media may also differ in age, country of origin or professional occupation (Kilian, Hennigs, & Langner, 2012; Murthy & Longwell, 2013).

Unlike other social media (for example, Facebook, YouTube, Instagram, Reddit), Twitter is often considered as the platform for news (Hermida, 2010). Most of the top world newspapers have their official accounts on Twitter platform (Orellana-Rodriguez & Keane, 2018). Besides, Twitter has millions of active users who can also generate their content. Lightning-fast reaction on various events demonstrated by Twitter users was confirmed by Sakaki, Okazaki, and Matsuo (2010). At the same time, for example, Bollen, Mao, and Zeng (2011) have shown that public mood of Twitter messages can predict changes in macroeconomic indicators (Dow Jones Industrial Average index closing values).

Having regard to the above, it is appropriate to consider Twitter data for EPU index calculation. Our study is not the first attempt to calculate EPU index based on Twitter data. Renault et al. (2021) have already calculated Twitter-based EPU index which was focused on only English-speaking part of data. The core problem of this index is that it does not reflect the whole picture for many countries where the core language is not English. In this paper, we account for this problem by considering native-speaking Twitter data for selected countries.

Our choice of countries for the analysis is based on several considerations. First, we consider only large enough countries since they provide data samples with enough number of bankrupt companies. Second, we choose countries from the same region (Europe) and, hence, they do not differ considerably in terms of culture. On one hand, this secures credibility of comparative analysis between the countries. On the other, such an approach reduces the bias in pooled estimations. Third, the role of core language (that is not English) in chosen countries must be dominant (over English) in business communications. This enables us to perform more reliable comparison of the effects of EPU indices computed based on articles in local newspapers (original EPU index) and tweets (our Twitter-based version of the index) which are both represented in native language in this study. Based on these reasoning we have selected four countries: France, Italy, Russia, and Spain.

The extent to which Twitter is used in the selected countries is also an important factor. According to Statista data (see Appendix A), by number of active users, France, Italy, and Spain are among the top Twitter users in the World while Twitter usage in Russia has been significantly lower. On one hand, this creates some heterogeneity in our sample with respect to Twitter usage. On the other, we can compare the

[1] Dixit and Pindyck (1994) review the earlier theoretical literature including studies by Oi (1961), Hartman (1972), and Abel (1983).

effectiveness of inclusion of original versus Twitter-based EPU indices in bankruptcy prediction models for countries with different level of Twitter usage.

Originally, the indices for different countries have been differently composed. For example, Baker et al. (2016) use different banks of words for different countries. The first bank of words is the same for all countries. It contained terms related to *Economy*, in particular, "economy" and "economic". The second bank of words is related to *Policy*. For the US the Policy bank includes the words/terms: "regulation", "deficit", "legislation", "congress", "white house", "Federal Reserve", "the Fed", "regulations", "regulatory", "deficits", "congressional", "legislative", and "legislature". It is easy to notice that Policy-related bank of words contains the terms which are inapplicable for other countries such as "congress", "white house", "Federal Reserve" and "the Fed". That is why they were changed by authors to their equivalents in other states. For example, for Russian index instead of "white house", the term "kremlin" is applied. The third bank of words is also the same in all countries. It stands for *Uncertainty* and includes words "uncertain" and "uncertainty".

Our Twitter dataset includes data for France, Italy, Russia, and Spain from 01.01.2014 until 31.12.2019. To compute Twitter-based index, we first downloaded all Twitter messages which contain a term from the first bank (*Economy*) *AND* a term from the second bank (*Policy*) in native language. Then we counted the number of messages in our dataset which contain terms from the third bank (*Uncertainty*) and computed their share from the total number of downloaded messages for the period (month). We got the final numbers with the same method as Baker et al. (2016) applied: we normalized obtained numbers to a mean of 100 for the entire time-series period. Because of the restrictions caused by accounting data which contain only year-to-year values, we had to convert monthly numbers to annual. We did it by computing simple average of monthly data. The total numbers of downloaded Twitter messages are reported in Appendix B.

3.1.1. Country specifics in calculation of Twitter-based EPU index

For index calculation we use only those messages which are written in the main language of each country. For such language as Italian, all messages are most probably related to Italy. For Russia the situation is similar. Even though there is a certain number of messages related to Ukraine, Belarus, Moldova and other former CIS/USSR countries, the events in these countries may cause the economic uncertainty in Russia (for example, Ukrainian crisis, unrest in Belarus). However, we cannot apply the same argument for France and Spain. Due to their colonial history, many French-speaking and Spanish-speaking people live outside these countries. That is why we need to identify only those messages which are related to France and Spain. Unfortunately, there is no effective way to detect the region of origin of Twitter message. To solve this issue, we applied the same method as Baker et al. (2016) used for Chinese index. We added one more country-related bank of words which contains country-oriented words. Hence, for France and Spain, all downloaded messages must contain a term from country-oriented bank of words. In Table 1 we report all words included in each bank of words for each country (in core/native language of the country).

We should note that we also calculated a modified version of the index which takes into consideration public uncertainties about economic crises. We included the term "crisis" and its derivatives into the original Uncertainty bank of words. Due to this modification, Twitter-based EPU index with crisis-related terms adds all messages which mention crisis-related topics into the group of posts which indicate uncertainty. One more advantage of our innovation is that it makes the index less volatile because of more messages which contain terms from the third bank of words. On Figs. 1–4 we depict the EPU indices for each country.

Table 1
Word lists.

Economic	*Policy*	*Uncertainty*	*Country*
France			
Economie, economique	taxe, impot, politique, regulation, reglementation, loi, depense, deficit, banque centrale, BCE, reserve, budget, budgetaire, monétaire, depense, cout, parlement, gouvernement, législation, taux, president, presidence, Élysée	incertitude, incertain, ambiguïté, inintelligibilité, imprecision + **crise**	France, français, française
Italy			
Economia, economie, economico, economici, economica, economiche	tassa, tasse, politica, politiche, regolamento, regolamenti, spesa, spese, deficit, Banca Centrale, Banca d'Italia, budget, bilancio	incerto, incerta, incerti, incerte, incertezza + **crisi**	NA
Russia			
эконоМика, эконоМический	политика, налог, расход, бюдЖет, расходование, регулирование, ЦБ РФ, дуМа, госдуМа, креМль, закон, законодательство, Монетарный, торговля, ставка	неопределённость, неопределённый, неизвестность, неясность, не ясность, не ясный, не ясен + **кризис**	NA
Spain			
económico, economía	impuesto, fiscal, monetario, tarifa, regulacion, reglamento, politica, gastar, gasto, presupuesto, presupuestario, presupuestal, deficit, banco central, congreso, cortes, ley, legislación, legislativo, legislador	incierto, incertidumbre, inseguridad + **crisis**	Spain, españa, español

Note: Not all forms of words are mentioned. In the real data all possible forms were applied.

3.2. Company level data

Bureau van Dijk Amadeus was used as a source of accounting data for the companies. We first considered company-level accounting information from 2014 to 2019 for France, Italy, Russia, and Spain. Phan, Sharma, and Tran (2018) and Yu, Fang, Du, and Yan (2017) studies have shown that construction and manufacturing industries are the most EPU-dependent areas. The same suggestions had the authors of the index themselves (Baker et al., 2016), who expected that EPU index should have particularly significant impact on policy dependent industries such as manufacturing and construction. Hence, our final data sample includes only construction and manufacturing companies. The other reason to choose these two sectors is that they are the most studied industries in bankruptcy topic.

Amadeus data contains a wide variety of indicators. All of them can be named as internal accounting-based variables. The full list of control internal (i.e., company-level) variables used in this study can be seen in Appendix C. Control internal indicators have been chosen based on the studies on bankruptcy topic. These measures have been applied by Platt et al. (1994), Dakovic, Czado, and Berg (2010), Amendola, Restaino, and

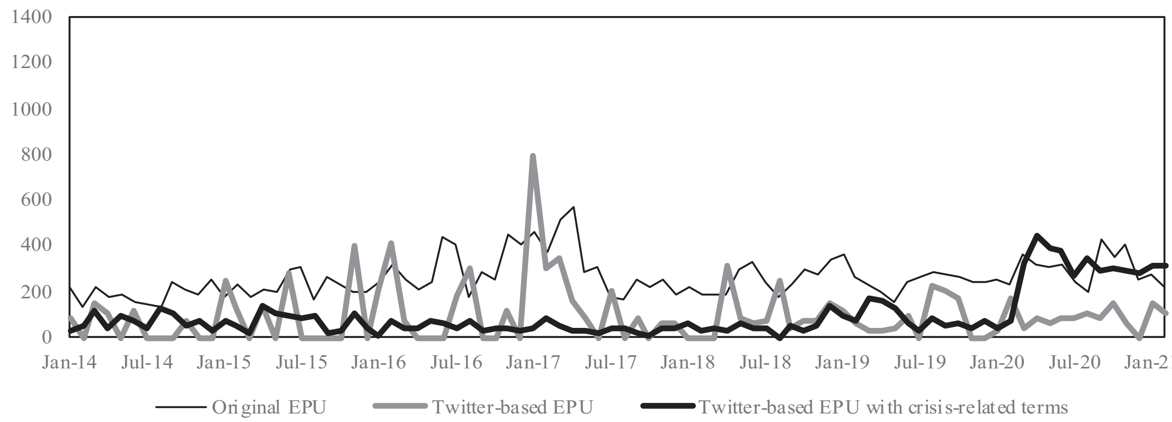

Fig. 1. EPU indices for France.

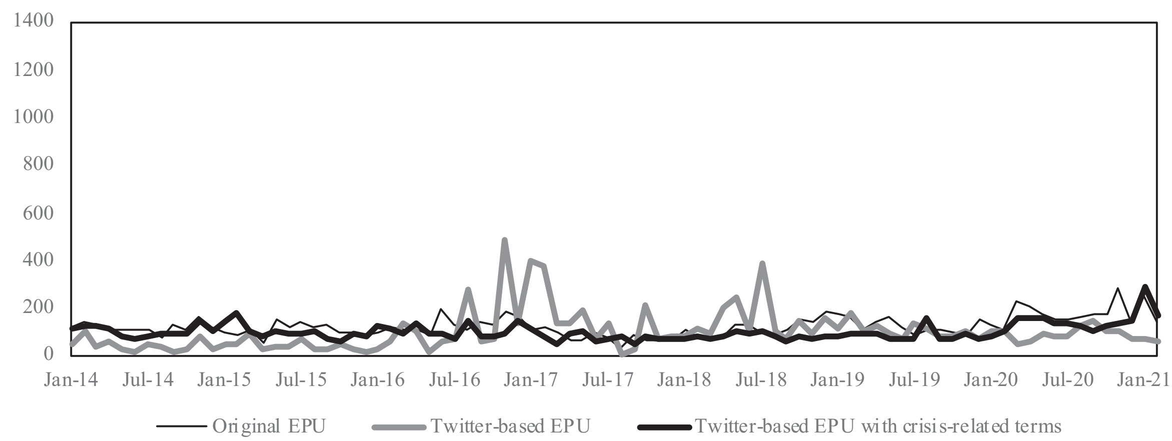

Fig. 2. EPU indices for Italy.

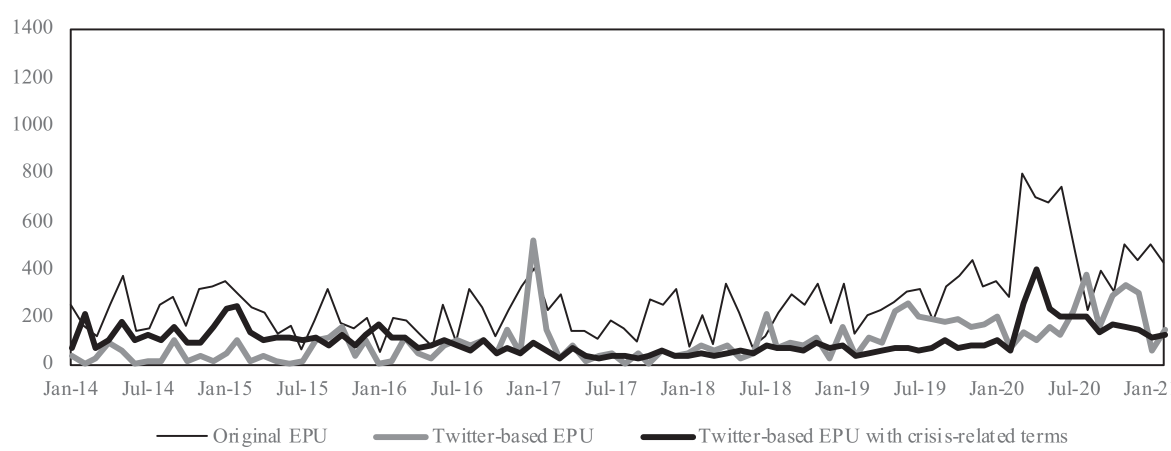

Fig. 3. EPU indices for Russia.

Sensini (2015), Du Jardin (2015) and others.

Initially our dataset contained information for 100,667 companies with more than 10 employees. It included both bankrupt and active companies. For bankrupt companies only one year before bankruptcy data was collected, while for active companies all possible data from 2014 to 2019 was included. Observations which contained blank values for considered variables have been removed from the dataset. After this, 300,553 annual observations remained. The number of bankrupt companies in the whole sample is 1001, 483 of which are Italian, 331 – Russian, 133 – Spanish and 54 – French.

We also control for several external (macro) factors. First, we include a 10 year government bond yield following Platt et al. (1994) who linked it with indebtedness of the company that plays an important role for bankruptcy. Following Santoro and Gaffeo paper (2009), we control for real GDP (PPP), inflation and real wage. Unemployment level, country's import and export indicators are also included as they can affect profitability measures (Dzikevičius & Šaranda, 2016). The description of variables and data sources are summarized in Appendix D. The

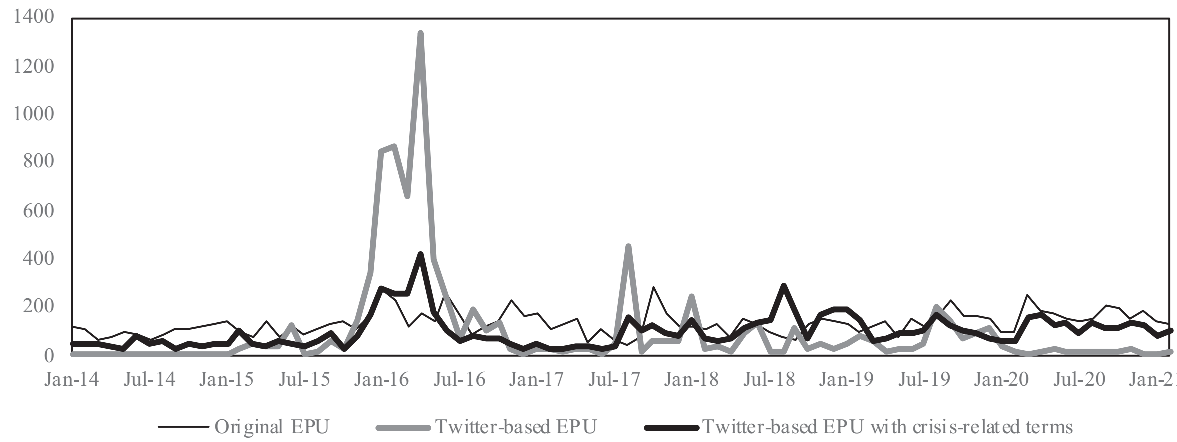

Fig. 4. EPU indices for Spain.

descriptive statistics of the variables is represented in Appendix E.

4. Method

4.1. Classification of models used in the study

There are two basic approaches to the bankruptcy prediction modeling based on machine learning (ML) methods: single and ensemble classifier models (see Table 2). Single classification techniques may involve statistical methods or supervised or unsupervised machine learning and rely on a single classifier. Previous literature has confirmed that particularly three single classifiers, namely, Support Vector Machine (SVM; see, for example, Min, Lee, & Han, 2006; Chaudhuri & De, 2011; Chen, 2011; Veganzones & Séverin, 2018), Logistic Regression (LR; see, for example, Jabeur, 2017; Son et al., 2019) and Artificial Neural Network (ANN; see, for example, Chen, Huang, & Lin, 2009; Nyitrai & Virág, 2019; Lee & Choi, 2013; Son et al., 2019) are the most effective in bankruptcy prediction models. Hence, we consider these three basic single classifiers in our baseline computations.

Second approach (ensemble classifier models) aims to combine several classifiers (i.e. classifiers of first level) to identify the most accurate classifier (see, for example, Polikar, 2012, Barboza et al., 2017). This approach has several advantages over single classifier models: higher stability to the absence of part of input variables, smaller variance (dispersion of algorithm responses due to the randomness of learning sample, noise in it and stochastic nature of settings) and bias (mathematical expectation of the difference between the true response and the response of the model, which characterizes the ability of the algorithm to approximate the objective function) (Geman, Bienenstock, & Doursat, 1992). Ensemble classifier models can be divided into three types: bagging, boosting, and stacking.

Bagging (or "bootstrap aggregating") is a technique involving independent classifiers that uses portions of the data and then combines them through model averaging, providing the most efficient results concerning a collection (Breiman, 1996). Random forests are used to perform this method (Breiman, 2001). Such ensembles usually perform better than single models (see, for example, Kim & Kang, 2010 and Choi, Son, & Kim, 2018) and depending on learning sample, this method can be very precise (Kruppa, Schwarz, Arminger, & Ziegler, 2013).

The boosting method consists of the repeated use of a base prediction rule or function on different sets of the initial set. Boosting builds on other classification schemes assigning a weight to each training set, which is then incorporated into the model (Begley, Ming, & Watts, 1996). The data are then reweighted (Barboza et al., 2017). Derivative algorithms as AdaBoost (adaptive boost) are successfully used for classification prediction (see, for example, Kim & Upneja, 2014).

Stacking (or stacked generalization) was proposed by Wolpert (1992) and aims to improve the classification performance of a single classifier by combining multiple classifiers in a two-level classification manner (Tsai & Hsu, 2013). In stacking, the outputs of individual classifiers are level zero generalizers that train the "stacked" classifier, which becomes a level one generalizer. The stacked classifier output determines the final decision. Hence, the stacked classifier is a trainable combiner that is different from ensembles classifiers (Tsai & Hsu, 2013).

From theoretical point of view, bagging algorithm is significantly easier to implement and every separate model provides rather high accuracy. Hence, technically, results` averaging cannot worsen the final result. However, stacking has a significant advantage over bagging as it automatically adjusts the recording of the results of various models (particularly of heterogenous models) that in turn secures increase in AUC. Recent research (see, for example, Ribeiro & dos Santos Coelho, 2020; Yang, Zhang, Lu, & Jin, 2021) also confirms that stacking ensembles outperform bagging algorithm approach.

4.2. Stages and research methods of the study

This study's research roadmap is presented on Fig. 5. In the first stage we prepare input datasets for the four cases: (a) without EPU index; (b) with original (newspaper based) EPU index; (c) with Twitter-based EPU index; (d) with Twitter-based EPU index with crisis-related terms. We first use datasets for all countries (France, Italy, Spain, and Russia) for each case. Next, we build models for companies from one country. However, the number of bankrupt companies in Spain and France is too small to get reliable one-country samples and, hence, one-country models have been built only for Italy and Russia.

To obtain reliable results, the data sample should be balanced. However, in our sample the share of bankrupt companies is rather small (unbalanced data is a common feature of bankruptcy studies). Hence, for test set we extracted 30% of bankrupt companies and added to them an equal number of active companies. Next, we have used two-step procedure to balance the data in the training set. First, we propagated the minority class (bankruptcy companies) several times that final sample would be rather big in size. After that we reduced the majority class

Table 2
Types of ML classifier models for the analysis of companies` bankruptcy.

Model		*Implementation*
Single classifier		Artificial neural network (ANN)
		Support vector machine (SVM)
		Logistic regression (LR)
Ensemble classifier	Bagging	Random forest (RF)
	Boosting	XGBoost
	Stacking	

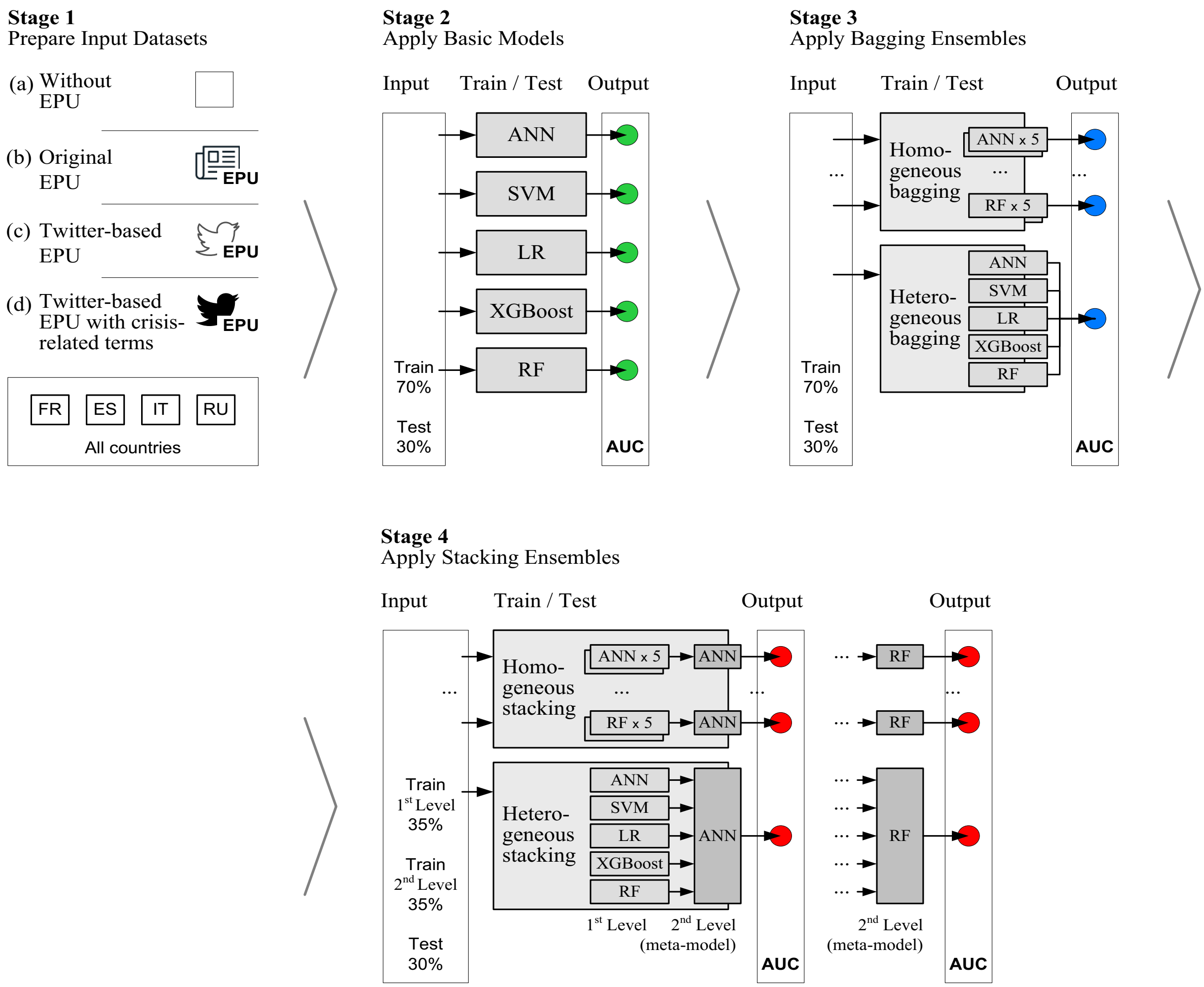

Fig. 5. Research roadmap.

(active companies) to the size of minority class by random removal (Veganzones & Séverin, 2018). Then we mixed up the sample in a random way. We also normalized the sample using z-normalization:

$$nD_i = \frac{D_i - \overline{D}}{Std(D)} \tag{1}$$

where nD_i – normalized data of indicator i; D_i – source value of indicator i; $\overline{D}$ – mean value of indicator i; $Std(Data)$ – standard deviation of indicator i. We should note that by using label encoding, we also included into the model nonnumeric indicator – industry label of a company.

After balancing the input data (that represents the first stage of the analysis), we evaluate the effectiveness of our ensembles. First, we apply basic classifier models (that represents the second stage of our analysis). We examine the preciseness of single models using Artificial Neural Network (ANN), Support Vector Machines (SVM) and Logistic Regression (LR). We also utilize evaluation methods of ready ensembles: XGBoost and Random Forest (RF). As metrics we use *Area Under the ROC Curve* (AUC).

On the third stage we apply the bagging ensembles. We chose weak learners, which reached highest accuracy on the first stage, and on their base built ensembles of five homogenous learners. We also tried an ensemble of heterogenous learners for all the models described above. For one set of models for the used dataset we utilized the same test set to verify all the models. Bagging ensemble algorithm is the following (see also Fig. 5). First, test data is downloaded for each weak learner to obtain probability of estimates. Second, based on the mean probability of estimates AUC is computed. Finally, for accuracy appraisal, the final classification, obtained by method of models` majority voting, is estimated.

Finally, on the fourth stage we apply the stacking ensembles. We first built ensembles of homogenous and heterogenous learners for first-level classifiers. To build the second-level classifier (meta-model) we used the same basic models as on the previous stage. Stacking ensemble algorithm is the following (see also Fig. 5). First, learning part of the sample is divided into two equal parts (that constitutes 35% of original number of bankruptcies, counting for separation of test set). This is needed because meta-model should learn on data that weak learners did not receive. Hence, first part of the learning sample is used to train weak learners on the first level. Their models are then saved, and their accuracy is estimated by test set that does not include copied exemplars of data. Next, weak learners get second part of the sample: each data exemplar is downloaded into weak learners. After that their responses are recorded as a row of new table that is used as a learning sample for all variants of meta-models. Similar approach was used in other studies (see, e.g., Jia et al., 2021). Meta-model reports the final ensemble response. Finally, the AUC is computed. To assess the advantage of stacking over bagging method, we build both ensembles on the same trained classifiers of the first level.

We utilize cross-validation (Stone, 1974) for basic models` training to obtain more reliable results. Cross-validation approach is based on dividing the sample into k parts and training models on different parts of the original sample. Cross-validation has various implementation schemes (Valente, Castellanos, Hausfeld, De Martino, & Formisano, 2021), however, to validate an ensemble, especially in conditions of an insufficiently large number of exemplars, we need a special approach. The previously described method of dividing and preparing the dataset is repeated 5 times to obtain several sets (training and test) of samples, thus reducing the impact of possible distortions due to a certain combination of active companies remaining after balancing the sample.

For single models` results` training we utilize these 5 subsets of datasets, each of which is tested on its training sample, while final accuracy metrics are computed as average for all 5 models. For bagging ensembles` validation we divide the training sample into 5 parts and each of the composite models is trained on different parts that is an important factor of their uniqueness in the ensemble. For the whole ensemble's reliability validation we repeat the procedure for each prepared dataset (5 in this study). Hence, we end up with 5 completely different ensembles, each of which is trained on its own variant of the balanced sample.

For the stacking ensemble, we perform similar actions to ensure the difference of the basic models due to 5-fold cross-validation. However, since in this case we need to divide the training dataset into two parts: for training the basic models and for training the meta-model, we pre-randomly divide 70% of the original dataset (the training part) into two parts of 35% of the original, and the use of 5 variants of the dataset also ensures the results reliability. All the results are the averages for the five variants of model construction.

The stages and machine learning (ML) classifier models used in this study are summarized in Table 3.

The parameters used in our models are summarized in Table 4.

4.3. Variables` selection

Random forests, trained on the whole sample, were used for variables` selection. According to prior research, this approach is effective for this task compared to alternative extraction of the needed characteristics from logistic regression (Aldrich & Auret, 2010). Results are reported in Table 5.

As we can see from the Table the level of significance of the variables can be clearly and expectedly divided into two levels: high level of significance for internal company-level indicators and low level of significance for external macro indicators. However, two internal indicators, CTA (cash/total assets) and LTD (long term debt) exhibit rather low level of significance and, hence, we have removed them from our data sample.

Table 3
Machine learning classifier models used in this study.

Basic models (Stage 2)	*Bagging ensembles (Stage 3)*	*Stacking ensembles (Stage 4)*
	Homogenous:	Homogenous:
ANN	ANN+ANN + ANN + ANN + ANN	(ANN + ANN + ANN + ANN + ANN) + ANN-meta
SVM	RF + RF + RF + RF + RF	(ANN + ANN + ANN + ANN + ANN) + RF-meta
LR		(ANN + ANN + ANN + ANN + ANN) + SVM-meta
XGBoost		(ANN + ANN + ANN + ANN + ANN) + LR-meta
RF		(RF + RF + RF + RF + RF) + ANN-meta
		(RF + RF + RF + RF + RF) + RF-meta
		(RF + RF + RF + RF + RF) + SVM-meta
		(RF + RF + RF + RF + RF) + LR-meta
	Heterogenous:	Heterogenous:
	ANN + SVM + LR + XGBoost+RF	(ANN + SVM + LR + XGBoost+RF) + ANN-meta
		(ANN + SVM + LR + XGBoost+RF) + RF-meta
		(ANN + SVM + LR + XGBoost+RF) + SVM-meta
		(ANN + SVM + LR + XGBoost+RF) + LR-meta

Table 4
Hyperparameter adjustments.

First level classifiers	*Parameter*	*Value*
NN	optimizer	adam
	epochs	100
	bath size	30
	hidden layers	2
	neurons	L1: 2048, L2: 1024
	save best only	True
LR	penalty	L2
	C	100
	Max_iter	100
SVM	kernel	rbf
	C	100
	gamma	auto
RF	n_estimators	100
	Max_depth	7
	criterion	gini
XGBoost	loss	deviance
	n_estimators	100
	learning_rate	0.1
	max_depth	3
Second-level classifiers	***Parameter***	***Value***
NN	optimizer	RMSprop
	epochs	30
	bath size	20
	hidden layers	1
	neurons	30
	save best only	True
LR	penalty	L2
	C	100
	Max_iter	100
SVM	kernel	rbf
	C	100
	gamma	auto
RF	n_estimators	50
	Max_depth	7
	criterion	gini

Table 5
The importance of variables based on Gini-coefficients.

№	Feature	Gini-importance	№	Feature	Gini-importance
High level:			14	CRP	0.045552
0	SLR	0.052724	15	Industry	0.045318
1	GE	0.050861	16	NWC	0.044240
2	Interest Cover	0.049322	17	NAT	0.043699
3	ROTA	0.048928	18	CAR	0.042504
4	QR	0.048707	**Low level:**		
5	ROCE	0.048605	19	CTA	0.033402
6	PR	0.047776	20	Import	0.012193
7	ROE	0.047479	21	Unempl	0.010657
8	COP	0.046686	22	GDP(PPP)	0.009342
9	Profit margin	0.046532	23	Infl	0.008603
10	ST	0.046404	24	Wage(real)	0.008322
11	FS	0.046370	25	Interest	0.008154
12	CR	0.046335	26	LTD	0.007948
13	ROA	0.045803	27	Export	0.007535

5. Results

5.1. Baseline results

Results of the second stage of the method (application of base classifier models) are reported in Table 6. We also tested four other models for single classifiers (Linear Discriminant Analysis, Stochastic Gradient Descent, Lasso and Ridge regressions), but their AUC values have been much lower and hence, we do not use them in our baseline estimations

Table 6
Results of basic models training (all countries).

Results	Without EPU					Original EPU				
	ANN	RF	XGBoost	SVM	LR	ANN	RF	XGBoost	SVM	LR
Training set AUC	**0,99543**	0,98203	0,97314	0,95147	0,87499	**0,99382**	0,97933	0,97058	0,95375	0,8814
Test set AUC	0,89287	0,9065	**0,91632**	0,90586	0,8488	0,90339	0,91053	**0,91314**	0,88539	0,84908
Mean avg.	0,94415	0,944265	**0,94473**	0,928665	0,861895	**0,94861**	0,94493	0,94186	0,91957	0,86524
Weighted mean avg.	**0,96466**	0,95937	0,95609	0,93779	0,86713	**0,96669**	0,95869	0,95335	0,93324	0,87170

Results	Twitter-based EPU					Twitter-based EPU with crisis-related terms				
	ANN	RF	XGBoost	SVM	LR	ANN	RF	XGBoost	SVM	LR
Training set AUC	**0,99656**	0,98036	0,9751	0,95676	0,88318	**0,99659**	0,98077	0,97094	0,95224	0,88388
Test set AUC	0,90291	**0,90683**	0,91974	0,90611	0,8476	0,91709	0,92124	**0,92813****	0,89636	0,86399
Mean avg.	**0,94974**	0,94360	0,94742	0,93144	0,86539	**0,95684**	0,95101	0,94954	0,92430	0,87394
Weighted mean avg.	**0,96847**	0,95830	0,95849	0,94157	0,87251	**0,97274***	0,96291	0,95810	0,93548	0,87791

Note: The best results for each EPU case (i.e. without EPU index and with various types of EPU index) are marked by bold. * denotes best result for the dataset among EPU cases (highest weighted mean average); ** denotes best result for test set among EPU cases (highest test set AUC).

(however, we report the respective results in Appendix F).

Based on AUC, we can evaluate the models` accuracy. Models with original EPU indices exhibit lower AUC than models with Twitter-based EPU indices. The highest AUC is observed for models with Twitter-based EPU indices adjusted for crisis.

From Table 6 we can evaluate the accuracy of training for base models and respective ensemble algorithms (RF and XGBoost) for the full sample. ANN and RF models exhibit the highest accuracy. Hence, we use them as a base to build other ensembles from homogenous learners (i.e. we apply two-step ensemble). We also verify the ensembles` accuracy when they are applied to heterogenous learners.

Results of the third stage (application of bagging ensembles) are reported in Table 7 – for all countries, in Table 8 – for Italy and in Table 9 – for Russia.

Several conclusions can be drawn from Tables 7–9. First, hetero bagging ensemble exhibits lowest AUC in all cases. The highest accuracy in training set is reached by ANN (in 6 from 12 cases; 4 cases are observed in all countries` dataset). The highest AUC in test set is achieved by RF (for all cases including separate datasets of Italian and Russian companies) that indicates its better ability to work with data not received during training. We can also conclude that RF is more effective than ANN in smaller datasets (i.e. country-specific). However, for the biggest dataset (which include all countries), RF has demonstrated the best ability to work with new data (test set). Finally, compared to base models in Table 6, the highest AUC (by weighted mean average; denoted by *) has increased from 0.97274 to 0.97464 (for the case of all countries).

Models with Twitter-based EPU index have highest AUC for datasets of all countries and Italy while for Russia models with original EPU index perform better than models with Twitter-based EPU indices. The plausible explanation is that Twitter is less popular in Russia than in European countries and, hence, Twitter-based EPU index is less informative.

Results of the fourth stage of the method (application of stacking ensembles) are reported in Table 10 – for all countries, in Table 11 – for Italy and in Table 12 – for Russia.

Results presented in Tables 10–12 indicate that ANN classifier of first level is not effective in stacking ensembles. Some RF ensembles perform better. Besides, by contrast with bagging, stacking ensemble of heterogenous models shows very good results (denoted by *Hetero* in tables). On the other hand, RF demonstrates ideal accuracy as a classifier for meta-model (AUC equals to one), though in test set it performs worse than other classifiers. ANN shows high results as meta-model though in many cases it is equivalent in accuracy or even performs worse than SVM and LR (e.g. for Italy the best model is hetero ensemble with LR meta-model).

The highest AUC by weighted mean average for the whole sample (i.e. that includes all countries) has increased from 0.97464 (bagging) to 0.97836 (stacking; corresponds to the model with RF classifier of the first level and RF meta-model). For test set, the highest AUC has increased from 0.93277 (bagging) to 0.93355 (stacking; corresponds to the model with RF classifier of the first level and ANN meta-model). Hence, though not all combinations of classifiers of the first level and meta-models outperform bagging, best combinations (for example, RF ensemble with ANN meta-model) outperform best variants of bagging. Consequently, we can conclude that stacking ensemble can build a classifier, which AUC will be higher than that of other models.

Models with Twitter-based EPU indices show best results as in previous stages (base classifier models and bagging). For all countries` sample, the highest accuracy has been achieved for the case with Twitter-based EPU with crisis-related terms. For Italy the best result has been reached with Twitter-based EPU index without crisis terms. For Russia the result somewhat different. Though the highest accuracy on test set was reached for the case with original EPU index (as in bagging), the best result by weighted mean average is for the case with Twitter-based EPU index without crisis terms. This can be due to measurement error because of roughly equal information capacity of both original and Twitter-based EPU indices in a relatively small dataset.

AUC results of all models used in this study are summarized on Fig. 6. On each plot we use Training set AUC on the abscissa and Test set AUC

Table 7
Results of bagging ensembles training (all countries).

Results	Without EPU			Original EPU			Twitter-based EPU		Twitter-based EPU with crisis-related terms			
	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hybrid	ANN	RF	Hetero
Training set AUC	**0,99731**	0,98323	0,97041	**0,99742**	0,98104	0,96893	**0,99798**	0,98166	0,97251	**0,99694**	0,98005	0,96808
Test set AUC	0,89968	**0,92238**	0,9094	0,90144	**0,92038**	0,9285	**0,91659**	0,91574	0,91751	0,92262	**0,93277****	0,92519
Mean avg.	0,94850	**0,95281**	0,93991	0,94943	**0,95071**	0,94872	**0,95729**	0,94870	0,94501	**0,95978**	0,95641	0,94664
Weighted mean avg.	**0,96802**	0,96498	0,95211	**0,96863**	0,96284	0,95680	**0,97356**	0,96188	0,95601	**0,97464***	0,96587	0,95521

Table 8
Results of bagging ensembles training (Italy).

Results	Without EPU			Original EPU			Twitter-based EPU			Twitter-based EPU with crisis-related terms		
	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero
Training set AUC	0,99326	**0,99625**	0,98396	0,99398	**0,9954**	0,98524	0,99541	**0,99607**	0,98278	**0,99697**	0,99593	0,98501
Test set AUC	0,92797	**0,93809**	0,94597	0,92816	**0,9335**	0,94769	0,94895	**0,95755****	0,9635	0,9352	**0,94077**	0,94476
Mean avg.	0,96061	**0,96717**	0,96496	0,96107	**0,96445**	0,96646	0,97218	**0,97681**	0,97314	0,96608	**0,96835**	0,96488
Weighted mean avg.	0,97367	**0,97880**	0,97256	0,97423	**0,97683**	0,97397	0,98147	**0,98451***	0,97699	0,97843	**0,97938**	0,97293

Table 9
Results of bagging ensembles training (Russia).

Results	Without EPU			Original EPU			Twitter-based EPU			Twitter-based EPU with crisis-related terms		
	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero
Training set AUC	0,9933	**0,99755**	0,99393	0,98726	**0,99787**	0,99258	**0,98925**	0,99762	0,99232	0,98835	**0,99799**	0,99166
Test set AUC	0,86992	**0,90633**	0,90727	0,88091	**0,92595****	0,91879	0,89015	**0,91273**	0,90992	0,89242	**0,91288**	0,91098
Mean avg.	0,93161	**0,95194**	0,9506	0,93408	**0,96191**	0,95568	0,9397	**0,95517**	0,95112	0,94038	**0,95543**	0,95132
Weighted mean avg.	0,95628	**0,97018**	0,96793	0,95535	**0,97629***	0,97044	0,95952	**0,97215**	0,9676	0,95957	**0,97245**	0,96745

Note for* Tables 7–9*: The best results for each EPU case (i.e. without EPU index and with various types of EPU index) are marked by bold. * denotes best result for the dataset among EPU cases (highest weighted mean average); ** denotes best result for test set among EPU cases (highest test set AUC).

Table 10
Results of stacking ensembles training (all countries).

Results	Without EPU			Original EPU		Twitter-based EPU			Twitter-based EPU with crisis-related terms			
	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero
ANN meta-model:												
Training set AUC	0,97501	0,95926	**0,97925**	0,9793	0,95955	**0,97987**	0,96924	0,95626	**0,97494**	0,97252	0,96011	**0,97568**
Test set AUC	0,8983	**0,90467**	0,89984	0,90133	0,92156	**0,92876**	**0,91653**	0,9075	0,90013	0,92114	**0,93355****	0,9284
Mean avg.	0,93666	0,93197	**0,93955**	0,94032	0,94056	**0,95432**	**0,94289**	0,93188	0,93754	0,94683	0,94683	**0,95204**
Weighted mean avg.	0,95200	0,94288	**0,95543**	0,95591	0,94815	**0,96454**	**0,95343**	0,94163	0,95250	0,95711	0,95214	**0,96150**
RF meta-model:												
Training set AUC	1	1	1	1	1	1	1	1	1	1	1	1
Test set AUC	0,87992	**0,89669**	0,84606	0,86797	**0,91113**	0,91027	0,88036	**0,89912**	0,89631	0,90297	**0,92785**	0,89869
Mean avg.	0,93996	**0,94835**	0,92303	0,93399	**0,95557**	0,95514	0,94018	**0,94956**	0,94816	0,95149	**0,96393**	0,94935
Weighted mean avg.	0,96398	**0,96901**	0,95382	0,96039	**0,97334**	0,97308	0,96411	**0,96974**	0,96889	0,97089	**0,97836***	0,96961
SVM meta-model:												
Training set AUC	0,96175	0,95925	**0,97902**	0,95782	0,94882	**0,97414**	0,95586	0,95653	**0,97333**	0,95011	0,96036	**0,97397**
Test set AUC	0,86579	**0,90482**	0,8906	0,89746	**0,91899**	0,9168	0,90067	**0,90762**	0,89371	0,90736	0,91297	**0,91848**
Mean avg.	0,91377	0,93203	**0,93481**	0,92764	0,933905	**0,94547**	0,92827	0,93208	**0,93352**	0,92874	0,93667	**0,94623**
Weighted mean avg.	0,93296	**0,94292**	0,95249	0,93971	0,93987	**0,95694**	0,95355	0,94191	**0,95311**	0,95767	0,95239	**0,96231**
LR meta-model:												
Training set AUC	0,97495	0,95928	**0,97908**	0,97945	0,95957	**0,97992**	0,96941	0,95659	**0,97473**	0,97234	0,96047	**0,97554**
Test set AUC	0,89841	0,90508	**0,90872**	0,90701	0,92149	**0,92759**	**0,91654**	0,90764	0,90267	0,92343	**0,93354**	0,93144
Mean avg.	0,93668	0,93218	**0,9439**	**0,94323**	0,94053	0,95375	**0,94298**	0,93212	0,93870	0,94789	0,94701	**0,95349**
Weighted mean avg.	0,95199	0,94302	**0,95797**	**0,95772**	0,94815	0,96422	**0,95355**	0,94191	0,95311	0,95767	0,95239	**0,96231**

Note: The best results for each EPU case (i.e. without EPU index and with various types of EPU index) are marked by bold. * denotes best result for the dataset among EPU cases (highest weighted mean average); ** denotes best result for test set among EPU cases (highest test set AUC).

on the ordinate. Colored dots represent AUC results for basic models (green), bagging ensembles (blue) and stacking ensembles (red).

The right upper quadrant represents zone with the highest accuracy (i.e. “ideal” zone). For all countries` sample, the model’s accuracy tends to increase from the case “without EPU index” to the case “with Twitter-based EPU index with crisis related terms”. The country-specific results (for Italian and Russian companies) are less clear though, on average, cases “with EPU indices” exhibit higher accuracy than the case “without EPU index”.

6. Conclusions

Prior research provided ample empirical evidence on the effects of economic policy uncertainty on macroeconomic dynamics and firm-level activities that in turn affect corporate bankruptcy decisions. In this study we suggest including economic policy uncertainty measure in bankruptcy prediction models directly. We utilize several machine learning techniques (in particular, single classifier models, bagging ensemble and stacking ensemble) to predict bankruptcy in the sample of French, Italian, Spanish, and Russian firms. We further compare the accuracy of different machine learning methods for bankruptcy prediction and show that stacking ensemble seems to outperform other

Table 11
Results of stacking ensembles training (Italy).

Results	Without EPU			Original EPU			Twitter-based EPU			Twitter-based EPU with crisis-related terms		
	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero
ANN meta-model:												
Training set AUC	0,96303	0,97055	**0,97408**	0,97283	0,97558	**0,9816**	0,97442	0,97972	**0,98607**	0,97174	0,97807	**0,9829**
Test set AUC	0,92867	**0,93832**	0,91604	0,92667	**0,93403**	0,90727	0,94695	**0,9586**	0,958	0,93622	**0,9427**	0,92508
Mean avg.	0,94585	**0,95443**	0,94506	0,94975	**0,95480**	0,94443	0,96068	0,96916	**0,97203**	0,95398	**0,96038**	0,95399
Weighted mean avg.	0,95272	**0,96088**	0,95666	0,95898	**0,96311**	0,95930	0,96617	0,97338	**0,97764**	0,96108	**0,96745**	0,96555
RF meta-model:												
Training set AUC	1	1	1	1	1	1	1	1	1	1	1	1
Test set AUC	**0,90979**	0,83979	0,89047	**0,91252**	0,81536	0,87909	**0,91657**	0,89606	0,90548	**0,9127**	0,85441	0,8449
Mean avg.	**0,95489**	0,91989	0,94523	**0,95626**	0,90768	0,93954	**0,95828**	0,94803	0,95274	**0,95635**	0,92720	0,9224
Weighted mean avg.	**0,97293**	0,95193	0,96714	**0,97375**	0,94460	0,96372	**0,97497**	0,96881	0,97164	**0,97381**	0,95632	0,9534
SVM meta-model:												
Training set AUC	0,94587	0,96964	**0,97384**	0,94944	0,97574	**0,98208**	0,95299	0,97965	**0,98551**	0,94535	0,97812	**0,98276**
Test set AUC	0,91207	**0,93758**	0,92699	0,9055	**0,93142**	0,91557	0,95012	**0,95422**	0,94821	0,92452	0,9103	**0,92751**
Mean avg.	0,92897	**0,95361**	0,95041	0,92747	**0,95358**	0,94882	0,95155	**0,96693**	0,96686	0,93493	0,94421	**0,95513**
Weighted mean avg.	0,93573	**0,96002**	0,95978	0,93625	**0,96244**	0,96212	0,95212	0,97202	**0,97432**	0,93910	0,95777	**0,96618**
LR meta-model:												
Training set AUC	0,96366	0,97086	**0,97356**	0,97273	0,97594	**0,98154**	0,97437	0,97977	**0,98567**	0,97302	0,97836	**0,98305**
Test set AUC	0,92881	**0,93855**	0,93758	0,92718	**0,93394**	0,93114	0,95026	0,80205	**0,96061****	0,93613	**0,94107**	0,93776
Mean avg.	0,94623	0,95470	**0,95557**	0,94995	0,95494	**0,95634**	0,96231	0,89091	**0,97314**	0,95457	0,95971	**0,96040**
Weighted mean avg.	0,95320	0,96116	**0,96276**	0,95906	0,96334	**0,96642**	0,96713	0,92645	**0,97815***	0,96195	0,96717	**0,96946**

Note: The best results for each EPU case (i.e. without EPU index and with various types of EPU index) are marked by bold. * denotes best result for the dataset among EPU cases (highest weighted mean average); ** denotes best result for test set among EPU cases (highest test set AUC).

Table 12
Results of stacking ensembles training (Russia).

Results	Without EPU			Original EPU			Twitter-based EPU			Twitter-based EPU with crisis-related terms		
	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero	ANN	RF	Hetero
ANN meta-model:												
Training set AUC	0,96453	0,98164	**0,9877**	0,95581	0,97819	**0,97888**	0,95887	0,97684	**0,97912**	0,94778	**0,9812**	0,98067
Test set AUC	0,86682	**0,90598**	0,85561	0,87489	**0,92523**	0,9197	0,89886	**0,91121**	0,88288	0,89326	**0,91352**	0,90519
Mean avg.	0,91567	**0,94381**	0,92165	0,91535	**0,95171**	0,94929	0,92886	**0,94402**	0,931	0,92052	**0,94736**	0,94293
Weighted mean avg.	0,93521	**0,95894**	0,94807	0,93153	**0,96230**	0,96112	0,94086	**0,95715**	0,95024	0,93142	**0,96089**	0,95802
RF meta-model:												
Training set AUC	1	1	1	1	1	1	1	1	1	1	1	1
Test set AUC	0,80379	0,80511	**0,83402**	0,87061	0,82898	**0,87534**	0,86114	0,87413	**0,88398**	0,88064	0,82364	0,84977
Mean avg.	0,90189	0,90255	**0,91701**	0,93530	0,91449	**0,93767**	0,93057	0,93706	**0,94199**	0,94032	0,91182	0,92488
Weighted mean avg.	0,94113	0,94153	**0,95020**	0,96118	0,94869	**0,96260**	0,95834	0,96223	**0,96519***	0,96419	0,94709	0,95493
SVM meta-model:												
Training set AUC	0,95333	0,98184	**0,98514**	0,93194	**0,97886**	**0,97659**	0,94329	**0,97677**	0,97635	0,93787	0,9811	**0,98141**
Test set AUC	**0,86386**	0,80894	0,85288	0,8725	**0,92561****	0,9222	0,86106	**0,91083**	0,89879	0,85939	0,86072	**0,89417**
Mean avg.	0,90859	0,89539	**0,91901**	0,90222	**0,95223**	0,94939	0,90217	**0,9438**	0,93757	0,89863	0,92091	**0,93779**
Weighted mean avg.	0,92648	0,92997	**0,94546**	0,91410	**0,96288**	0,96027	0,91862	**0,95698**	0,95308	0,91432	0,94498	**0,95523**
LR meta-model:												
Training set AUC	0,96506	0,98181	**0,98552**	0,95591	**0,97898**	0,97889	0,95695	0,97691	**0,97825**	0,94847	0,98107	**0,98081**
Test set AUC	0,86682	**0,90614**	0,89068	0,88098	**0,92545**	0,92341	0,8978	**0,91144**	0,89841	0,89379	0,9061	**0,91083**
Mean avg.	0,91594	**0,94397**	0,9381	0,91844	**0,95221**	0,95115	0,92737	**0,94417**	0,93833	0,92113	0,94358	**0,94582**
Weighted mean avg.	0,93558	**0,95910**	0,95706	0,93343	**0,96292**	0,96224	0,93920	**0,95726**	0,95429	0,93206	0,95857	**0,95981**

Note: The best results for each EPU case (i.e. without EPU index and with various types of EPU index) are marked by bold. * denotes best result for the dataset among EPU cases (highest weighted mean average); ** denotes best result for test set among EPU cases (highest test set AUC).

methods, though marginally.

To measure economic policy uncertainty, we utilize a modification of Baker et al. (2016) index. We compute a Twitter-based version of the index considering the tweets in native language. Our analysis provides firm and robust evidence that bankruptcy prediction models with EPU index always outperform models without EPU index irrespective the method (single classifier, bagging or stacking ensemble) and the type of the index (original EPU or its Twitter-based modification) used. We further demonstrate that, on average, our Twitter-based modification of the EPU index performs better in bankruptcy prediction than the original index.

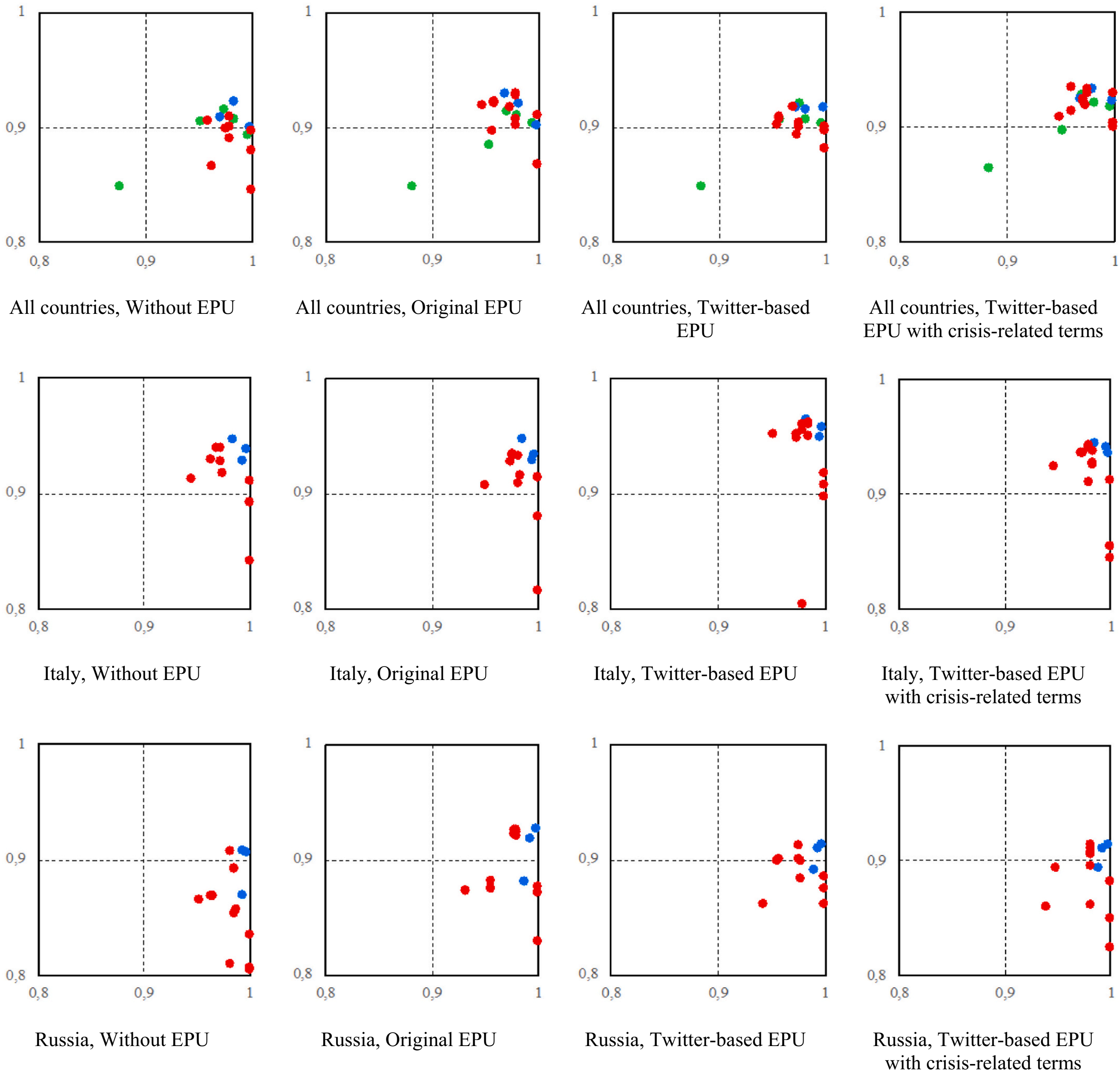

Fig. 6. AUC results.

Data availability

Data will be made available on request.

Acknowledgement

We would like to thank Dmitriy Afanasyev, Igor Demin, Vera Kononova, Yuri Zelenkov and webinar participants for helpful comments on earlier drafts of this paper. We also thank Artiom Cozin for excellent research assistance. During part of the research for this publication Svetlana Ledyaeva was hosted by the Aleksanteri Institute, University of Helsinki.

Appendix A

Table A1
Twitter usage in selected OECD and BRICS countries.

Country	Twitter users, millions	Date of statistics	% from population of 2020
Japan	58.2	October 2021	46.25
United Kingdom	19.05	October 2021	28.34
USA	77.75	October 2021	23.60
Canada	8	October 2021	21.05
Spain	9.35	October 2021	19.75
Turkey	16.25	October 2021	19.27
Italy	11.2	March 2021	18.81
Netherlands	2.9	2021	16.63
France	10.2	October 2021	15.14
Korea	7.2	October 2021	13.90
Chile	2.25	January 2021	11.77
Mexico	14	October 2021	10.86
Germany	7.8	October 2021	9.37
Brazil	19.05	October 2021	8.96
Colombia	4.3	October 2021	8.45
Costa Rica	0.39	January 2021	7.66
Russia	9	January 2021	6.25
India	24.45	October 2021	1.77

Sources: Authors` computations based on Statista data.

Appendix B

Table B1
Number of Twitter messages (total; downloaded).

	2014	2015	2016	2017	2018	2019	**SUM**
France	10216	7562	8650	11540	21880	28691	88539
Italy	49108	32057	26987	25367	60866	78375	272760
Russia	98402	93341	62325	56845	49132	44028	404073
Spain	13217	11747	12619	7738	19884	22657	87862
SUM	170943	144707	110581	101490	151762	173751	853234

Appendix C

Table C1
List of accounting/internal indicators.

Tag	Indicator	Formula
PM	Profit margin	PM = (PBT / OR) × 100
EBIT	Earnings before interest and tax	EBIT = R − E − DA
ROE	Return on equity	ROE = EA / SE
ROCE	Return on capital employed	ROCE = EBIT/Capital Employed
ROA	Return on assets	ROA = NI / TA
ROTA	Return on total assets	ROTA = (NI + IE + T) / TNA
NAT	Net assets turnover	NAT = OR / (SF + NCL)
IC	Interest cover	IC = OP / IP
ST	Stock turnover	ST = OR / STO
COP	Collection period	COP = (D / OR) × 360
CRP	Credit period	CRP = (C / OR) × 360
CR	Current ratio	CR = CA / CL
QR	Quick ratio	QR = (CA − STO) / CL
SLR	Shareholders liquidity ratio	SLR = SF / NCL
PR	Proprietary ratio	PR = (SF / TA) × 100
GE	Gearing	GE = ((NCL + LO) / SF) × 100
FS	Firm size	Log of TA
CAR	Cash asset ratio	CAR = C / CL
NWC	Net working capial	NWC=CA-CL
CTA	Cash/total assets	CTA = C/TA
LTS	Long term debt/shareholders funds	LTS = LTD / SF

Note: C = Creditors; CA = Current assets; CE = Cash & Equivalents; CL = Current liabilities; D = Debtors; DA = Depreciation & Amortization; EA = Earnings; E = Expense; IE = Interest Expense; IP = Interest paid; LO = Loans; NI = Net Income; NCL = Non-current liabilities; OP = Operating profit; OR = Operating revenue; PBT = Profit before tax; R = Revenue; SE = Shareholders Equity; SF = Shareholders funds; STO = Stock; T = Taxes; TA = Total Assets; TNA = Total Net Assets.

Appendix D

Table D1
List of external indicators.

Indicator	Description	Who applied	Data source
Prime interest rate	10-Year Bond Yield	Platt, H., Platt, M., & Pedersen, J. (1994). Bankruptcy discrimination with real variables. *Journal of Business Finance & Accounting*, 21(4), 491–510.	Investing.com
Real GDP (PPP,$)	Real GDP adjusted for the Purchasing Power Parity value represented in US Dollars	Santoro, E., & Gaffeo, E. (2009). Business failures, macroeconomic risk and the effect of recessions on long-run growth: A panel cointegration approach. *Journal of Economics and Business*, 435–452.	World bank
Inflation	Inflation, consumer prices (annual %)	Santoro, E., & Gaffeo, E. (2009). Business failures, macroeconomic risk and the effect of recessions on long-run growth: A panel cointegration approach. *Journal of Economics and Business*, 435–452.	World bank
Real wage	Nominal wage divided by Consumer Price Index represented in Euro	Santoro, E., & Gaffeo, E. (2009). Business failures, macroeconomic risk and the effect of recessions on long-run growth: A panel cointegration approach. *Journal of Economics and Business*, 435–452.	International Labour Organization
Unemployment level	Unemployment, total (% of total labour force) (modeled ILO estimate)	Dzikevicius, A., & Saranda, S. (2016). Establishing a set of macroeconomic factors explaining variation over time of performance in business sectors. *Business: Theory and Practice*, 17(2), 159–166.	World bank
Export	Exports, Million US dollars	Dzikevicius, A., & Saranda, S. (2016). Establishing a set of macroeconomic factors explaining variation over time of performance in business sectors. *Business: Theory and Practice*, 17(2), 159–166.	OECD
Import	Imports, Million US dollars	Dzikevicius, A., & Saranda, S. (2016). Establishing a set of macroeconomic factors explaining variation over time of performance in business sectors. *Business: Theory and Practice*, 17(2), 159–166.	OECD

Appendix E

Table E1
Descriptive statistics.

	Mean	Standard deviation	Min	Max	Skewness	Kurtosis
Profit margin	−9.03	1789.16	−769125.00	16372.02	−308.66	119617.74
ROE	0.15	6.44	−492.03	114.46	−129.17	64078.23
ROCE	0.07	4.89	−300.73	77.07	−337.86	160012.05
ROA	−0.15	1.90	−70.82	3.74	−525.20	283652.59
ROTA	−0.06	0.72	−27.17	1.85	−529.19	286704.55
NAT	4.80	61.83	−22435.00	10989.67	−130.42	62493.31
Interest cover	75.87	1232.89	−227711.00	303641.00	62.70	21338.28
ST	58.91	444.36	−0.02	18651.00	174.45	43039.26
COP	363.31	3785.94	−1094.21	120402.00	378.75	157315.81
CRP	551.61	7129.15	−1122.63	286146.00	279.69	86119.36
CR	2.30	18.57	0.01	522.05	389.39	171099.07
QR	1.76	17.22	0.00	502.33	393.50	172872.01
SLR	0.46	478.04	−14159.00	22251.00	136.21	28757.66
PR	0.00	1.92	−70.82	0.97	−147.37	48035.92
GE	1051.03	32727.92	−91138.10	1409600.00	183.66	48624.86
FS	8.62	1.54	2.48	17.35	0.45	0.56
CAR	0.19	0.76	0.00	29.47	501.71	264944.82
NWC	860.08	90283.65	−6002000.00	1184939.00	85.12	15779.21
CTA	0.06	0.09	0.00	0.94	2.06	5.03
LTS	7.97	300.25	−792.10	12953.50	−220.23	113786.42
Interest	2.74	2.60	0.13	11.36	1.86	2.31
Real GDP (PPP)	2.60E+12	7.32E+11	1.56E+12	4.43E+12	0.81	0.15
Inflation	1.46	2.64	−0.50	15.53	3.50	14.38
Wage(real)	1939.69	630.74	434.58	2849.53	−1.33	0.85
UnemplOUMENT LEVEL	12.09	4.94	4.60	24.44	0.80	0.34
Export	752105.73	171226.64	522555.00	1157399.00	0.93	−0.08
Import	675833.87	141338.72	471866.00	971674.00	0.58	−0.78
EPU (original)	156.56	64.67	78.22	317.12	1.03	0.00
EPU (Twitter-based)	105.10	79.05	2.89	406.44	2.40	6.96
EPU (Twitter-Based with crisis)	86.48	28.40	39.48	154.39	0.56	−0.07

Appendix F

Table F1
Results of basic models training (all countries).

Result	Without EPU				Original EPU			
	LDA	SGD	Lasso	Ridge	LDA	SGD	Lasso	Ridge
Training set AUC	0.76886	0.87301	0.72603	0.79723	0.86293	0.87723	0.79374	0.82973
Test set AUC	0.84157	0.87064	0.70412	0.79475	0.85980	0.87589	0.78955	0.82315

(*continued on next page*)

Table F1 (*continued*)

Result	Without EPU				Original EPU			
	LDA	SGD	Lasso	Ridge	LDA	SGD	Lasso	Ridge
Mean avg.	0.80522	0.87183	0.71508	0.79599	0.86137	0.87656	0.79165	0.82644
Weighted mean avg.	0.79067	0.87230	0.71946	0.79649	0.86199	0.87683	0.79248	0.82776

Result	Twitter-based EPU				Twitter-based EPU with crisis-related terms			
	LDA	SGD	Lasso	Ridge	LDA	SGD	Lasso	Ridge
Training set AUC	0.85205	0.82699	0.73021	0.80366	0.85615	0.86801	0.79633	0.81126
Test set AUC	0.84158	0.82674	0.72148	0.79752	0.85260	0.86761	0.79290	0.79352
Mean avg.	0.84682	0.82687	0.72585	0.80059	0.85438	0.86781	0.79462	0.80239
Weighted mean avg.	0.84891	0.82692	0.72759	0.80182	0.85509	0.86789	0.79530	0.80594

LDA – Linear Discriminant Analysis, SGD – Stochastic Gradient Descent, Lasso and Ridge – linear regression types

LDA – Linear Discriminant Analysis, SGD – Stochastic Gradient Descent, Lasso and Ridge – linear regression types.

References

Abel, A. B. (1983). Optimal investment under uncertainty. *The American Economic Review, 73*(1), 228–233.

Abid, A. (2020). Economic policy uncertainty and exchange rates in emerging markets: Short and long runs evidence. *Finance Research Letters, 37*, 101378.

Aldrich, C., & Auret, L. (2010). Fault detection and diagnosis with random forest feature extraction and variable importance methods. *IFAC Proceedings Volumes, 43*(9), 79–86.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609.

Amendola, A., Restaino, M., & Sensini, L. (2015). An analysis of the determinants of financial distress in Italy: A competing risks approach. *International Review of Economics & Finance, 37*, 33–41.

Antonakakis, N., Chatziantoniou, I., & Filis, G. (2014). Dynamic spillovers of oil price shocks and economic policy uncertainty. *Energy Economics, 44*, 433–447.

Ashraf, B. N., & Shen, Y. (2019). Economic policy uncertainty and banks' loan pricing. *Journal of Financial Stability, 44*, 100695.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics, 131*(4), 1593–1636.

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 83*, 405–417.

Basu, S., & Bundick, B. (2012). *Uncertainty shocks in a model of effective demand*. No. w18420.

Beaver, W. H. (1966). Financial ratios as predictors of failure journal of accounting research. *Empirical Research in Accounting: Selected Studies, 4*, 71–111.

Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies, 1* (4), 267–284.

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial education*, 1–42.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics, 98*(1), 85–106.

Bhattacharjee, A., Higson, C., Holly, S., & Kattuman, P. (2009). Macroeconomic instability and business exit: Determinants of failures and acquisitions of UK firms. *Economica, 76*(301), 108–131.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8.

Bordo, M. D., Duca, J. V., & Koch, C. (2016). Economic policy uncertainty and the credit channel: Aggregate and bank level US evidence over several decades. *Journal of Financial Stability, 26*, 90–106.

Breiman, L. (1996). Bagging predictors. *Machine Learning, 24*(2), 123–140.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Brogaard, J., & Detzel, A. (2015). The asset-pricing implications of government economic policy uncertainty. *Management Science, 61*(1), 3–18.

Caggiano, G., Castelnuovo, E., & Figueres, J. M. (2017). Economic policy uncertainty and unemployment in the United States: A nonlinear approach. *Economics Letters, 151*, 31–34.

Cepni, O., Guney, I. E., & Swanson, N. R. (2020). Forecasting and nowcasting emerging market GDP growth rates: The role of latent global economic policy uncertainty and macroeconomic data surprise factors. *Journal of Forecasting, 39*(1), 18–36.

Chaudhuri, A., & De, K. (2011). Fuzzy support vector machine for bankruptcy prediction. *Applied Soft Computing, 11*(2), 2472–2486.

Chen, H. J., Huang, S. Y., & Lin, C. S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. *Expert Systems with Applications, 36*(4), 7710–7720.

Chen, M. Y. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers & Mathematics with Applications, 62*(12), 4514–4524.

Chen, P. F., Lee, C. C., & Zeng, J. H. (2019). Economic policy uncertainty and firm investment: Evidence from the US market. *Applied Economics, 51*(31), 3423–3435.

Choi, H., Son, H., & Kim, C. (2018). Predicting financial distress of contractors in the construction industry using ensemble learning. *Expert Systems with Applications, 110*, 1–10.

Dakovic, R., Czado, C., & Berg, D. (2010). Bankruptcy prediction in Norway: A comparison study. *Applied Economics Letters, 17*(17), 1739–1746.

Demir, E., & Ersan, O. (2017). Economic policy uncertainty and cash holdings: Evidence from BRIC countries. *Emerging Markets Review, 33*, 189–200.

Dewaelheyns, N., & Van Hulle, C. (2008). Legal reform and aggregate small and micro business bankruptcy rates: Evidence from the 1997 Belgian bankruptcy code. *Small Business Economics, 31*(4), 409–424.

Dixit, R. K., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton University Press.

Drobetz, W., El Ghoul, S., Guedhami, O., & Janzen, M. (2018). Policy uncertainty, investment, and the cost of capital. *Journal of Financial Stability, 39*, 28–45.

Du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research, 242*(1), 286–303.

Dunham, L. M., & Garcia, J. (2021). Measuring the effect of investor sentiment on financial distress. *Managerial Finance, 47*(12), 1834–1852.

Dzikevičius, A., & Šaranda, S. (2016). Establishing a set of macroeconomic factors explaining variation over time of performance in business sectors. *Verslas: teorija ir praktika, 17*(2), 159–166.

Fitzpatrick, P. J. (1932). *A comparison of the ratios of successful industrial enterprises with those of failed companies*.

Friedman, M. (1968). The role of monetary policy. *American Economic Review, 58*, 1–17.

García-Perdomo, V., Salaverría, R., Kilgo, D. K., & Harlow, S. (2018). To share or not to share: The influence of news values and topics on popular social media content in the United States, Brazil, and Argentina. *Journalism Studies, 19*(8), 1180–1201.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation, 4*(1), 1–58.

Gilchrist, S., Sim, J. W., & Zakrajšek, E. (2014). *Uncertainty, financial frictions, and investment dynamics*. w20038. National Bureau of Economic Research.

Gulen, H., & Ion, M. (2016). Policy uncertainty and corporate investment. *The Review of Financial Studies, 29*(3), 523–564.

Hartman, R. (1972). The effects of price and cost uncertainty on investment. *Journal of Economic Theory, 5*(2), 258–266.

Hassett, K. A., & Metcalf, G. E. (1999). Investment with uncertain tax policy: Does random tax policy discourage investment. *The Economic Journal, 109*(457), 372–393.

Hermida, A. (2010). Twittering the news: The emergence of ambient journalism. *Journalism Practice, 4*(3), 297–308.

Higgs, R. (1997). Regime uncertainty: Why the great depression lasted so long and why prosperity resumed after the war. *The Independent Review, 1*(4), 561–590.

Hol, S. (2007). The influence of the business cycle on bankruptcy probability. *International Transactions in Operational Research, 14*(1), 75–90.

Hou, F., Tang, W., Wang, H., & Xiong, H. (2021). Economic policy uncertainty, marketization level and firm-level inefficient investment: Evidence from Chinese listed firms in energy and power industries. *Energy Economics, 100*, 105353.

Iqbal, U., Gan, C., & Nadeem, M. (2020). Economic policy uncertainty and firm performance. *Applied Economics Letters, 27*(10), 765–770.

Istrefi, K., & Piloiu, A. (2014). *Economic policy uncertainty and inflation expectations*. Banque de France Working Paper No. 511.

Jabeur, S., Stef, N., & Carmona, P. (2022). Bankruptcy prediction using the XGBoost algorithm and variable importance feature engineering. *Computational Economics*, 1–27.

Jabeur, S. B. (2017). Bankruptcy prediction using partial least squares logistic regression. *Journal of Retailing and Consumer Services, 36*, 197–202.

Jia, R., Lv, Y., Wang, G., Carranza, E., Chen, Y., Wei, C., & Zhang, Z. (2021). A stacking methodology of machine learning for 3D geological modeling with geological-geophysical datasets, Laochang Sn camp, Gejiu (China). *Computers & Geosciences, 151*, 104754.

Kang, W., Lee, K., & Ratti, R. A. (2014). Economic policy uncertainty and firm-level investment. *Journal of Macroeconomics, 39*, 42–53.

Kilian, T., Hennigs, N., & Langner, S. (2012). Do millennials read books or blogs? Introducing a media usage typology of the internet generation. *Journal of Consumer Marketing., 29*(12), 114–124.

Kim, M. J., & Kang, D. K. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert Systems with Applications, 37*(4), 3373–3379.

Kim, S. Y. (2018). Predicting hospitality financial distress with ensemble models: The case of US hotels, restaurants, and amusement and recreation. *Service Business, 12*(3), 483–503.

Kim, S. Y., & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and AdaBoosted decision tree models. *Economic Modelling, 36*, 354–362.

Kruppa, J., Schwarz, A., Arminger, G., & Ziegler, A. (2013). Consumer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications, 40*(13), 5125–5131.

Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques–a review. *European Journal of Operational Research, 180* (1), 1–28.

Laitinen, E. K., & Laitinen, T. (1998). Cash management behavior and failure prediction. *Journal of Business Finance & Accounting, 25*(7–8), 893–919.

Leduc, S., & Liu, Z. (2016). Uncertainty shocks are aggregate demand shocks. *Journal of Monetary Economics, 82*, 20–35.

Lee, S., & Choi, W. S. (2013). A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications, 40*(8), 2941–2946.

Levy, A., & Bar-Niv, R. (1987). Macroeconomic aspects of firm bankruptcy analysis. *Journal of Macroeconomics, 9*(3), 407–415.

Li, X. (2020). The impact of economic policy uncertainty on insider trades: A cross-country analysis. *Journal of Business Research, 119*, 41–57.

Liang, D., Tsai, C. F., Dai, A. J., & Eberle, W. (2018). A novel classifier ensemble approach for financial distress prediction. *Knowledge and Information Systems, 54*(2), 437–462.

Liang, D., Tsai, C. F., Lu, H. Y. R., & Chang, L. S. (2020). Combining corporate governance indicators with stacking ensembles for financial distress prediction. *Journal of Business Research, 120*, 137–146.

Liu, G., & Zhang, C. (2020). Economic policy uncertainty and firms' investment and financing decisions in China. *China Economic Review, 63*, 101279.

Liu, L., & Zhang, T. (2015). Economic policy uncertainty and stock market volatility. *Finance Research Letters, 15*, 99–105.

Lyandres, E., & Zhdanov, A. (2013). Investment opportunities and bankruptcy prediction. *Journal of Financial Markets, 16*(3), 439–476.

Min, S. H., Lee, J., & Han, I. (2006). Hybrid genetic algorithms and support vector machines for bankruptcy prediction. *Expert Systems with Applications, 31*(3), 652–660.

Murthy, D., & Longwell, S. A. (2013). Twitter and disasters: The uses of Twitter during the 2010 Pakistan floods. *Information, Communication & Society, 16*(6), 837–855.

Nam, C. W., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting, 27*(6), 493–506.

Nguyen, C. P., Le, T. H., & Su, T. D. (2020). Economic policy uncertainty and credit growth: Evidence from a global sample. *Research in International Business and Finance, 51*, 101118.

Ninh, B. P. V., Do Thanh, T., & Hong, D. V. (2018). Financial distress and bankruptcy prediction: An appropriate model for listed firms in Vietnam. *Economic Systems, 42* (4), 616–624.

Nyitrai, T., & Virág, M. (2019). The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences, 67*, 34–42.

Oi, W. Y. (1961). The desirability of price instability under perfect competition. *Econometrica: Journal of the Econometric Society*, 58–64.

O'Leary, D. E. (1998). Using neural networks to predict corporate failure. *Intelligent Systems in Accounting, Finance & Management, 7*(3), 187–197.

Orellana-Rodriguez, C., & Keane, M. T. (2018). Attention to news and its dissemination on Twitter: A survey. *Computer Science Review, 29*, 74–94.

Panousi, V., & Papanikolaou, D. (2012). Investment, idiosyncratic risk, and ownership. *The Journal of Finance, 67*(3), 1113–1148.

Pástor, Ľ., & Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics, 110*(3), 520–545.

Phan, D. H. B., Iyke, B. N., Sharma, S. S., & Affandi, Y. (2021). Economic policy uncertainty and financial stability–Is there a relation? *Economic Modelling, 94*, 1018–1029.

Phan, D. H. B., Sharma, S. S., & Tran, V. T. (2018). Can economic policy uncertainty predict stock returns? Global evidence. *Journal of International Financial Markets, Institutions and Money, 55*, 134–150.

Phan, H. V., Nguyen, N. H., Nguyen, H. T., & Hegde, S. (2019). Policy uncertainty and firm cash holdings. *Journal of Business Research, 95*, 71–82.

Pisula, T. (2020). An ensemble classifier-based scoring model for predicting bankruptcy of polish companies in the Podkarpackie Voivodeship. *Journal of Risk and Financial Management, 13*(2), 37.

Platt, H. D., Platt, M. B., & Pedersen, J. G. (1994). Bankruptcy discrimination with real variables. *Journal of Business Finance & Accounting, 21*(4), 491–510.

Polikar, R. (2012). Ensemble learning. In *Ensemble machine learning* (pp. 1–34). Boston, MA: Springer.

Putra, S. G. P., Joshi, B., Redi, J., & Bozzon, A. (2020 June). A credit scoring model for smes based on social media data. In *International conference on web engineering* (pp. 113–129). Cham: Springer.

Qu, Y., Quan, P., Lei, M., & Shi, Y. (2019). Review of bankruptcy prediction using machine learning and deep learning techniques. *Procedia Computer Science, 162*, 895–899.

Renault, T., Baker, S. R., Bloom, N., & Davis, S. (2021). Twitter-based uncertainty indices. Retrieved from Economic policy uncertainty: https://www.policyuncertainty.com/twitter_uncert.html.

Ribeiro, M. H. D. M., & dos Santos Coelho, L. (2020). Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Applied Soft Computing, 86*, 105837.

Rodrik, D. (1991). Policy uncertainty and private investment in developing countries. *Journal of Development Economics, 36*(2), 229–242.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010 April). Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web* (pp. 851–860).

Santoro, E., & Gaffeo, E. (2009). Business failures, macroeconomic risk and the effect of recessions on long-run growth: A panel cointegration approach. *Journal of Economics and Business, 61*(6), 435–452.

Sarikov, B., & Kuprianov, A. (2020). The effect of monetary policy on corporate bankruptcies: Evidence from the United States. *Applied Economics, 52*(47), 5158–5168.

Son, H., Hyun, C., Phan, D., & Hwang, H. J. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications, 138*, 112816.

Stolbov, M., & Shchepeleva, M. (2020). Systemic risk, economic policy uncertainty and firm bankruptcies: Evidence from multivariate causal inference. *Research in International Business and Finance, 52*, 101172.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological), 36*(2), 111–133.

Tsai, C. F., & Hsu, Y. F. (2013). A meta-learning framework for bankruptcy prediction. *Journal of Forecasting, 32*(2), 167–179.

Valente, G., Castellanos, A. L., Hausfeld, L., De Martino, F., & Formisano, E. (2021). Cross-validation and permutations in MVPA: Validity of permutation strategies and power of cross-validation schemes. *NeuroImage*, 118145.

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems, 112*, 111–124.

Wang, J., Yang, J., Iverson, B. C., & Kluender, R. (2020). *Bankruptcy and the COVID-19 crisis*. Available at SSRN 3690398.

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks, 5*(2), 241–259.

Yang, M., Zhang, J., Lu, H., & Jin, J. (2021). Regularized ELM bagging model for tropical cyclone tracks prediction in South China Sea. *Cognitive Systems Research, 65*, 50–59.

Yu, H., Fang, L., Du, D., & Yan, P. (2017). How EPU drives long-term industry beta. *Finance Research Letters, 22*, 249–258.

Zavgren, C. V. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic analysis. *Journal of Business Finance & Accounting, 12*(1), 19–45.

Zhang, G., Han, J., Pan, Z., & Huang, H. (2015). Economic policy uncertainty and capital structure choice: Evidence from China. *Economic Systems, 39*(3), 439–457.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E. P., Yan, H., & Li, X. (2011 April). Comparing twitter and traditional media using topic models. In *European conference on information retrieval* (pp. 338–349). Berlin, Heidelberg: Springer.