Ratio Mathematica Volume 43, 2022

An adaptive Neural Network approach to predict the Capital Adequacy Ratio

Giacomo Di Tollo*
Gerarda Fattoruso†
Bartolomeo Toffano‡

Abstract

Financial institutions, policy makers and regulatory authorities need to implement stress tests in order to test both resilience and the consequences of adverse shocks. The European Central Bank and the European Banking Authority regularly conduct these tests, whose importance is more and more evident after the financial crisis of 2007-2008. The stress tests' non-linear features of variables and scenarios triggered the need of general and robust strategies to perform this task. In this paper we want to introduce an adaptive Neural Network approach to predict the Capital Adequacy Ratio (CAR), which is one of the main ratios monitored to retrieve useful information along many stress test procedures. The Neural Network approach is based on a comparison between feed-forward and recurrent networks, and is run after a meaningful pre-processing operations definition. Results show that our approach is able to successfully predict CAR by using both Neural Networks and recurrent networks.

Keywords: Capital Adequacy Ratio; Stress Tests, Neural Network Approach.[1]

*Department of Law, Economics, Management and Quantitative Methods (DEMM), University of Sannio, Benevento, Italy; giditollo@unisannio.it

†Corresponding Author. Department of Law, Economics, Management and Quantitative Methods (DEMM), University of Sannio, Benevento and NEOMA BS, Rouen, France; Italy; fattoruso@unisannio.it

‡Department of Economics, Ca' Foscari University, Venice, Italy; bartolomeo.toffano@unive.it

[1]Received on July 20, 2022. Accepted on September 20, 2022. Published on September 25, 2022. doi: 10.23755/rm.v43i0.841. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

Banks' bankruptcy may have catastrophic effects over the overall economy, since the contagion effect it may trigger could lead to a generalised overall crisis [47]. To this extent the activity of banking supervision, and the role and the authority of bank regulation, play a big role in preventing (or reducing the effect of) the banks' bankruptcy [11, 44]. Although aimed to different targets, many of these supervision exercises are designed to maintain a sufficient banks' level of capital adequacy to allocate specific reserves aimed to face expected losses and to protect themselves again excessive credit expansion. Authorities regulations impose constraints over these reserves, even though banks often imposes themselves reserves higher than the ones imposed by the regulations. These constraints are defined by a minimum capital adequacy ratio that measures the level of capital as a function of the risk bore the bank [27].

In this framework, *systemic risk* represents the risk of breakdown of the entire financial system: this can be triggered by the misbehavior of a single component of the overall financial system, that triggers negative impacts on the overall system. This scenario can be observed on the timeline of years 2007-2008, starting from some *cracks* in the subprime mortgage markets leading to a worldwide financial crisis: this confirmed the idea that the more complex and non-linear a system, the higher the probability of the system to fail [23, 28]. In order to prevent these failures (and/or to quantify financial (un)stability), financial institutions resort to *stress tests*, which are non linear tools used to assess the magnitude of an exogenous shock and to determine a collapse threshold. These tools operate by investigating both the local stress level (i.e. a single bank) and how the shock is globally spread. Back in the 1990s, stress tests were intended for testing the *resilience* and the stability of a financial institution to reach a certain credibility. Later, stress tests, have been used to check the stability and vulnerability of single financial institutions and the overall banking systems [16]. Based on the evaluation aim and on the implications of the findings, stress tests can be classified into two major groups: microprudential stress tests, which are forward-looking supervisory instruments for determining the liquidity adequacy of individual banks in relation to their portfolio risks [4]; macroprudential stress tests, which consists of two different types of approaches: the bottom-up and the top-down approach [16]. In the bottom-up approach, the effect is measured using data on individual portfolios. On the other hand, in the top-down approach, the impact is estimated by using aggregated data. Many computational methods have been introduced to perform stress test and to predict bank failures: Discriminant Analysis [37], Logit and Probit analysis [8], Neural Networks [53], just to name a few. There are also contributions that compared different methods: for instance, [2] investigates different methods such as Logistic Regression (LR), Linear Discriminant

Analysis (LDA), Random Forests (RF), Support Vector Machines (SVM), Neural Networks (NN) and Random Forests of Conditional Inference Trees (CRF); [18] compared Generalized Linear Models and Generalized Additive Models, and concluded that Generalized Linear Mixed Models have a better ability to predict troubled businesses.

During a stress-test procedure many variables are taken into account to monitor the financial institutions, and there exist several studies aimed to assess the relative *importance* of these variables, in order to select what variable to monitor to get the most accurate information as possible. Many contributions focuses in determining the relationships between Capital Ratios and bank failures [1], and to understand whether these ratios are useful to assess the regulatory capital adequacy [42]. In this context, a careful investigation of Capital Ratios is crucial for both the regulatory authority and the bank itself, since it has been shown that the familiar banking characteristics for identifying a distress-prone bank identified fragile banks effectively during the global crisis without new information and are likely to continue to work well in the future [40]. Amongst many variables (e.g., profitability, liquidity, solvency, productivity, asset quality, see [54]), Capital Adequacy Ratio (CAR) has a prominent place, and has been used in the predictor set by many works [39, 31, 14, 38, 33, 5, 48]. Recently, some contributions proposed to predict it as an indicator of financial health [49]. In our contribution we want to expand this framework by implementing an adaptive Neural Network approach to predict CAR from a well established set of indicators, and to provide banks and regulators with useful information about their stress-test activity.

Our contribution is organised as follows: Section 2 reports the main literature on the topic; Sections 3 and 4 outline the set of data and the pre-processing operation performed on it; Section 5 introduces the methods used in this paper; Section 6 comments the main results and Section 7 concludes the paper.

2 Literature review

CAR represents a particularly relevant topic for assessing the risks to which banks are exposed [6]. In fact, for the construction of the CAR index, credit risk, market risk, interest rate risk and exchange rate risk are considered. In this sense, the regulatory authorities define the CAR as a significant indicator of safety and stability as it considers capital as a useful element to absorb losses [34]. Currently, the Capital Adequacy Ratios (CARs) defined by the minimum ratio of capital to risk weighted assets are 8% under Basel II and 10.5% under Basel III [3, 12]. Based on this, CAR represents a factor of analysis by regulators to determine capital adequacy for banks and to perform stress tests [29]. In order to aggregate the information coming from the literature, we performed an analysis on bibliographic

data using the software *VOSviewer* (Figure 1) to create a keyword co-occurrence map in order to analyze the main CAR literature in our field of analysis.

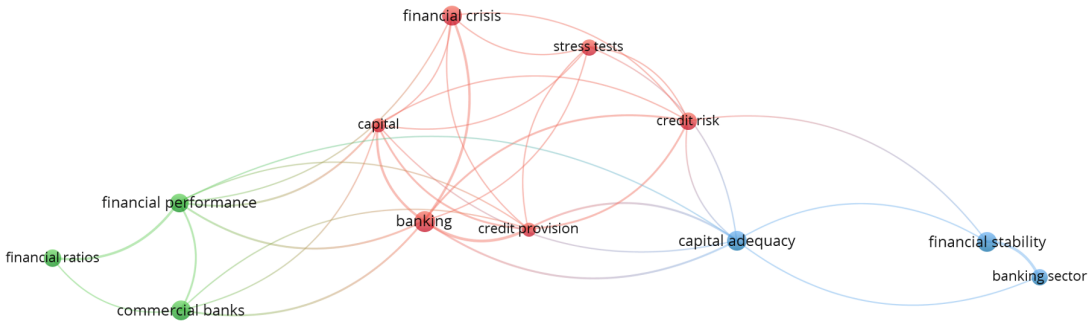

Figure 1: VOSviewer: Capital Adequacy Ratio

From the analysis of the data, it emerges that several authors analyze the required minimum levels of the CAR by evaluating macroeconomic indicators [46], [7], financial indicators [50], multi credit rating indicators [45]. Furthermore, many authors carry out stress tests on CAR to verify the effects of economic crises [58], stability [25] and resilience [15] of banks, along with macro stress test for resilience assessment [20]. Recent studies are moving towards identifying the most important variables for future projections of the CAR. In particular, [50] carry out a study on South Korean national banks using Random Forest Boruta algorithms, Random Forest Recursive Feature Elimination, and Bayesian Regularization Neural Networks. Other contributions use CAR to benchmark the performances of banks in stress tests [3, 12, 24, 27, 29, 32, 59].

The goal of our contribution is to assess whether we can use stress-testing to effectively benchmark the performance of a bank in a precise scenario, and to this extent we need to choose a metrics that can precisely fill that role. CAR is apt to measures the financial soundness of banks in absorbing a reasonable amount of loss, and on the basis of the central role that the CAR assumes in the assessments of banks and on the basis of the guidelines of the literature on the analysis of the minimum levels of the CAR, our work aims to accurately predict the CAR by using quantitative methods. In this framework the quantitative research about stress-testing has been twofold: on one side, to predict the banks' bankruptcy; on the other side, to assess the different variables features and capability to explain the default. According to [54], Neural Networks are widely used in contributions related to the first side, while its application about the other side are still limited. We can start our discussion by pointing out that along with stress testing, a key topic is the prediction of various risks, that was based on traditional probability

and statistical theories [9], but that could lead to non-linear formulations or to taking into account just a few variables, hence triggering the we need of complex and non linear models, also due to the needs of a more interconnected world, not only in financial terms. For this reason researchers and risk managers avail themselves of the usage of Artificial Neural Networks and Deep Learning to stress testing activities and predict high volatility periods. Since more hidden layers are in a Neural Network means a more complex modelling interaction effect, in finance forecasts, large collections of data often require dynamic data relationships that are difficult or impossible to specify under a complete model [9]. On the other hand, deep learning models can identify and manipulate dynamic non-linear data connections that are invisible to any current financial economic theory and may deliver more reliable predictive outcomes than traditional approaches [30]. Although Artificial Neural Networks and Deep Learning have several applications in the financial field such as credit scoring, predictions and forecasts in financial crisis and bankruptcy, we want to focus on how Artificial Neural Networks and Deep Learning Methods are related to stress testing. Financial stability is essential to the economic growth of countries and individuals. Regulatory agencies and foreign organisations carried out stress testing activities to determine the stability of the financial system even earlier than 2007, but failed to anticipate the unprecedented economic implications of the crisis. For this reason ever more stress testing exercises were created and used from the authorities with a glance to the consequences of an interconnected financial system in the macroeconomic environment. For example, the European Banking Authority (namely, EBA) approach uses simplified assumptions that cover only particular risks to individual bank balance sheets depending on the macro-economic scenario. One of the major drawbacks of the European Banking Authority approach is the static financial statement expectation, which allows assets and liabilities to stay stable over the horizon considered without any appreciation of management decisions or new loans. Macroeconomic feedback impacts, such as the influence of large insolvent firms on the global economy, are not generally welcomed assumptions in these systems. This kind of test aims the planning binaries of an after crisis recovery behaviour. However, [52] show that the main problem with respect to the European Banking Authority approach is that this mechanism does not provide an early alarm to avoid being completely disarmed in front of a shock. In [52] it is also provided a solution to this weakness of the model. They propose a Neural intelligence for which financial or macroeconomic disturbances extend to the bank's balance sheets while simultaneously building a large Neural Network with macro and financial factors. The model is capable of gathering more knowledge concealed in a large data set and allows for complex non-linear interactions that materialize under adverse macroeconomic conditions and financial strain. This methodology examines the financial system independently, without relying on the

forecasts of the single banks. As a result of the cited paper, comparing the static stress test models with dynamic ones, prove that the deep learning framework can become a useful tool and can improve the early warning mechanism's signaling ability to anticipate future financial issues and failures of individual banks. The authors, finally, compare the performance of the Deep Learning technique with the classic stress test models, such as the constant balance sheet approach and the dynamic balance sheet approach to satellite modelling. They reveal that the prediction error of the CAR dropped significantly under the Deep Learning Method due to its improved performance in simulating the one-year gains and losses of financial institutions. For this reason, Deep Learning Architecture may become a useful tool for macro prudential stress testing and can improve the early warning mechanism's ability to anticipate future financial crises and failures of individual banks.
Now that we have a measure by which we can benchmark banks, we need to find a way by which we can predict the CAR of banks based on certain factors which is what we need in order to stress test banks. The more factors we can incorporate in our predictions the better since it will reflect better a real-world situation and make our stress testing much more realistic.

3 Data set

Our data set consists of worldwide banks' financial indicators; along with stress financial indicators, we have considered also macro-economic indicators, in order to identify the propagation of systemic shocks that propagate into the financial institutions. We have retrieved quarterly observations that covers a period of 12 years (2007 to 2019).

Data was collected from different sources: stress financial indicators have been collected from the Federal Deposit Insurance Corporation[2] website[3]; macroeconomics indicators were collected from the Federal Reserve Economic Data[4] website[5]. The sample period covers twelve years: we have collected quarterly data referring to 672 banks and financial institutions between 2007 and 2019, hence we dispose of 34944 observations. The sample includes missing and noisy

[2] *FDIC is an independent agency created by the Congress to maintain stability and public confidence in the nation's (USA) financial system. The FDIC insures deposits; examines and supervises financial institutions for safety, soundness, and consumer protection; makes large and complex financial institutions resolvable; and manages receiverships*

[3] *referred to as https://www.fdic.gov/FDIC in what follows*

[4] *referred to as https://fred.stlouisfed.org/FRED in what follows*

[5] *Researchers at the St. Louis Fed contribute to monetary policy discussions by advising on a range of topics, especially in preparation for Federal Open Market Committee meetings* (from the https://www.fdic.gov/FDIC website, accessed on 2021, January 29^{th}).

Table 1: Variables used in the experimental phase and the category they belong to: the label *FIN* denotes financial indicators and *MAC* denotes macro-economic indicators.

Name	Description	*FIN / MAC*
net loan	Net loans and leases exposure	FIN
loss allow	Loss allowance to loans	FIN
dep	Total deposits	FIN
yield ea	Yield on earning assets	FIN
fundc ea	cost of funding earning assets	FIN
inc aa	Noninterest income to average assets	FIN
CAR	Total risk-based capital ratio	FIN
tot asst	Average total assets	FIN
tot eq	Average total equity	FIN
tot loan	Average total loans	FIN
risk dens	Risk weight density	FIN
GDP growth	Gross Domestic Product growth	MAC
export growth	US real exports of goods and services growth	MAC
debt GDP	US public debt to GDP	MAC
govex GDP	US government expenditure to GDP	MAC
inflat	Implicit price deflator as a measure of US inflation	MAC
HPI growth	House Price Index growth	MAC
unemp	Unemployment rate (age 15-64)	MAC
Yield 10Y	10-year US sovereign bonds yields	MAC
SP500 ret	SP 500 quarterly returns	MAC

values. Please notice that we have chosen a sample period that does not contain sub-periods denoting the emergence of a crisis, since we want to develop a methodology for *ordinary* periods, in which systemic shocks are more difficult to detect. Collected data show a number of correct entries which is smaller than its theoretical value: this is due to missing and noisy data, and could lead to misbehavior of the Neural Network approach, hence we had to devise pre-processing operations, that are outlined in what follows.

4 Data pre-processing

Data analysis is a key point in all experimental settings, and it is always performed in order to understand its features, to detect anomalies (if any), and to represent data without loosing useful information. Based on the observations oulined by [13, 21], we apply the following data pre-processing operations.

Table 2: Variables used in the experimental phase: overall main statistics before pre-processing operations.

Name	Mean	STD	Kurt.	Skewn.	Min	Max
net loan	684062.10	2845178	128.86	10.17	0	75190000
loss allow	10811.40	92866.23	1170.43	29.11	0	5752000
dep	848100.51	4672760	498.48	18.76	68	2180000
yield ea	4.61	1.21	20.01	2.38	0.07	26.96
fundc ea	1.24	0.88	2.67	1.32	0	16.59
inc aa	1.47	18.11	722.73	26.40	-15.95	601.27
CAR	23.68	15.42	113.18	6.08	0.75	725.80
tot asst	1114221	5587407	302.58	14.46	2816	2110361
tot eq	125720.10	566772.30	129.17	10.33	539.75	14389800
tot loan	683913.22	2852402	129.57	10.19	0	71201027
risk dens	60.09	14.28	0.63	0.06	8.43	192.24
GDP growth	1.73	2.35	4.05	-1.53	-8.45	5.51
exp. growth	3.63	8.17	5.14	-1.19	-28.65	25.84
debt GDP	93.95	11.27	1.38	-1.49	61.65	105.18
govex GDP	0.34	0.01	-1.09	0.37	0.31	0.37
inflat	100.43	4.51	-0.94	0.12	91.70	111.25
HPI growth	201.26	20.39	0.22	0.92	176.86	264.31
unemp	7.22	1.89	-1.40	-0.07	3.78	10.05
Yield 10Y	2.64	0.72	-0.02	0.70	1.56	4.84
SP500 ret	0.02	0.06	6.76	-1.47	-0.27	0.17

4.1 Removal and replacement

When collecting data, one may incur in missing and incorrect values. Previous contributions related to Neural Network approaches [21] suggested to remove indicators containing more than 30% of missing and wrong values. Our set of data does not contain such indicators, so we are using the whole set of variables in our experimental phase. Anyhow, many indicators show missing and wrong values, so we replace missing values (due to computational errors) with the upper limit of the normalization (see what follows), and wrong values with the indicator's average over time.

4.2 Normalization

Normalization is a general procedure performed in order to feed the Neural Network with data belonging to the same range: many contributions stress the importance of performing meaningful normalization, and many formulas are sug-

gested [35]. In our case we used the logarithmic transformation that has been already introduced by [21], defined as follows:

$$\overline{x}_i = \log_u \left(|\min(0, x_{min})| + x_i + 1 \right), \tag{1}$$

where x_i represents the value before normalisation of input x for firm i, and $\overline{x}_i$ represents its normalised value. Please notice that we have defined

$$u$$

such that

$$u = x_{\text{max}} + 1$$

this has been imposed in order to have $\overline{x}_i \in [0, 1]$.

Table 3: Main statistics of overall financial and macro-economic indicators after pre-processing operations.

Name	Mean	STD	Kurtosis	Skewness	Min	Max
net loan	0.67	0.08	1.41	0.42	0	1
loss allow	0.48	0.10	1.29	0.25	0	1
dep	0.64	0.07	1.38	0.58	0.22	1
yield ea	0.53	0.06	5.19	-0.07	0.02	1
fundc ea	0.30	0.14	-0.28	0.53	0	1
inc aa	0.41	0.06	6.68	0.05	0	1
CAR	0.50	0.07	1.23	0.86	0.08	1
tot asst	0.65	0.07	1.38	0.66	0.42	1
tot eq	0.63	0.08	1.36	0.71	0.38	1
tot loan	0.67	0.08	1.43	0.43	0	1
risk dens	0.79	0.05	1.48	-0.74	0.44	1
GDP growth	1.06	0.31	2.58	-1.52	0	1
export growth	0.99	0.2	10.45	-3.09	0	1
debt GDP	0.97	0.02	2.11	-1.71	0.89	1
govex GDP	0.93	0.03	-1.09	0.36	0.86	1
inflat	0.98	¡ 0.01	-1.17	0.03	0.96	1
HPI growth	0.96	0.01	-0.83	0.01	0.92	1
unemp	0.86	0.10	-1.27	-0.28	0.65	1
Yield 10Y	0.79	0.11	-0.93	0.10	0.55	1
SP500 ret	1.12	0.57	-1.08	-0.09	0	1

4.3 Correlation analysis

We have performed a correlation analysis in order to understand whether some kind of correlation arise amongst variables defined in Section 3 and to avoid feeding the network with highly-correlated indicators. We have tested Pearson's, Kendall's, and Spearman's correlation, leading to similar trends in the obtained correlations. In what follows we will refer to Spearman's ranked based correlation. We have decided to remove from the predictor set the indicators showing a correlation with a given portion (i.e., j) of other indicators greater than a given threshold (i.e., h). In order to determine the value of j and h we have defined parameter-tuning procedure via REVAC (see [43]). The values found have been $j = \frac{1}{3}$ and $h = 0.70$. On the basis of these values, we have decided to remove indicators showing a correlation with 30% of the other indicators greater than 0.7. These indicators are: net_{loan}, $loss_{allow}$, dep, tot_{asst}, tot_{eq}, tot_{loan}. They will not be considered in what follows: 6 indicators have been removed from the predictors set, corresponding to 24 quarterly indicators that will not be used to feed the Neural Networks' nodes.

As for the Neural Network experiments, in what follows we are outlining results of the experiments run by using data before the pre-processing operations (referred to as *full model*) and data after the pre-processing operations (referred to as *reduced model*).

5 Experimental analysis

In this section we are introducing the methods used to perform our experimental analysis: the Neural Network approach will be detailed in Section 5.1, along with the main components needed to define its use, i.e., the network topologies (Section 5.1.1) and the partitioning of the set of data to enforce generalisation (Section 5.1.2). Then, we are introducing the methods we are comparing our approach with (Linear Regression in Section 5.2, and Generalised Linear Models in Section 5.3), along with the metrics used for our comparisons in Section 5.4.

5.1 Neural networks

In this section we are introducing our Neural Network approach: Artificial Neural Networks [22] can be referred to as algorithms that mimic the behavior of the human brain to perform complex tasks, and they are used to grasp non-functional relationships over the data. They are composed of elementary units (*neurons*) which are connected to each other via weighted and oriented links (*synapses*). Neurons may have different functions: the *input* neurons receive data

from external sources; the *output* neurons show the computed output values; the *hidden* neurons are used to perform computations. During the *learning* phase, the weights associated to the synapses iteratively change over time, accordingly to a specific algorithm: several algorithms have been proposed for this *learning* phase: *Back-Propagation* [56], *Quasi-Newton methods* [55], *Levenberg-Marquardt algorithm* [26], just to name a few. The *learning* phase may be organised following three different paradigms: *supervised* learning [36]; *unsupervised* learning [10] and *reinforcement* leaning [57]. In what follows, we are training networks introduced in Section 5.1.1 by using *back-propagation* algorithm, in order to minimize the network test set's root mean square error (RMSE) defined as

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - a_i)^2}, \tag{2}$$

where n is the test set size, e_i is expected output value corresponding to pattern i, and a_i is the actual network output corresponding to input pattern i.

5.1.1 Network topologies

For our experiments we are using two different Neural topologies: a *feed-forward*[6] architecture with 80 inputs nodes, referred to as *standard* network (see Figure 2), and a variant in which inputs neurons corresponding to the same indicator are grouped by 4 before feeding the first feed-forward layer[7] (see Figure 3).

Please notice that, as for the cardinality of hidden neurons, many rules of thumb exist, suggesting different formulas to compute the number of hidden layers and the number of hidden neurons [19]. We have decided not to use any of these rules, resorting to an adaptive method to determine the optimal hidden neurons structure: this procedure has been proposed by [17], and it aims to minimize the network's error (in our scenario, the Eq. 2) calculated for each of the data set at hand. This adaptive procedure starts with a single hidden neuron and iteratively add one neuron until no improvement on the Eq. 2 is found over the last user-defined K iterations, and is outlined in Algorithm 1.

[6] A feed-forward network features neurons grouped into layers (1,2, ..., l_{max}) : each neuron belonging to layer i ($i < l_{max}$) is associated to synapses that connect itself to all neurons belonging to layer $i + 1$.

[7] These four values correspond to past observations spreading over one year, since for each indicator i corresponding to time t the input pattern contains the value of the indicators collected at time t, together with the 3 previously quarterly collected values.

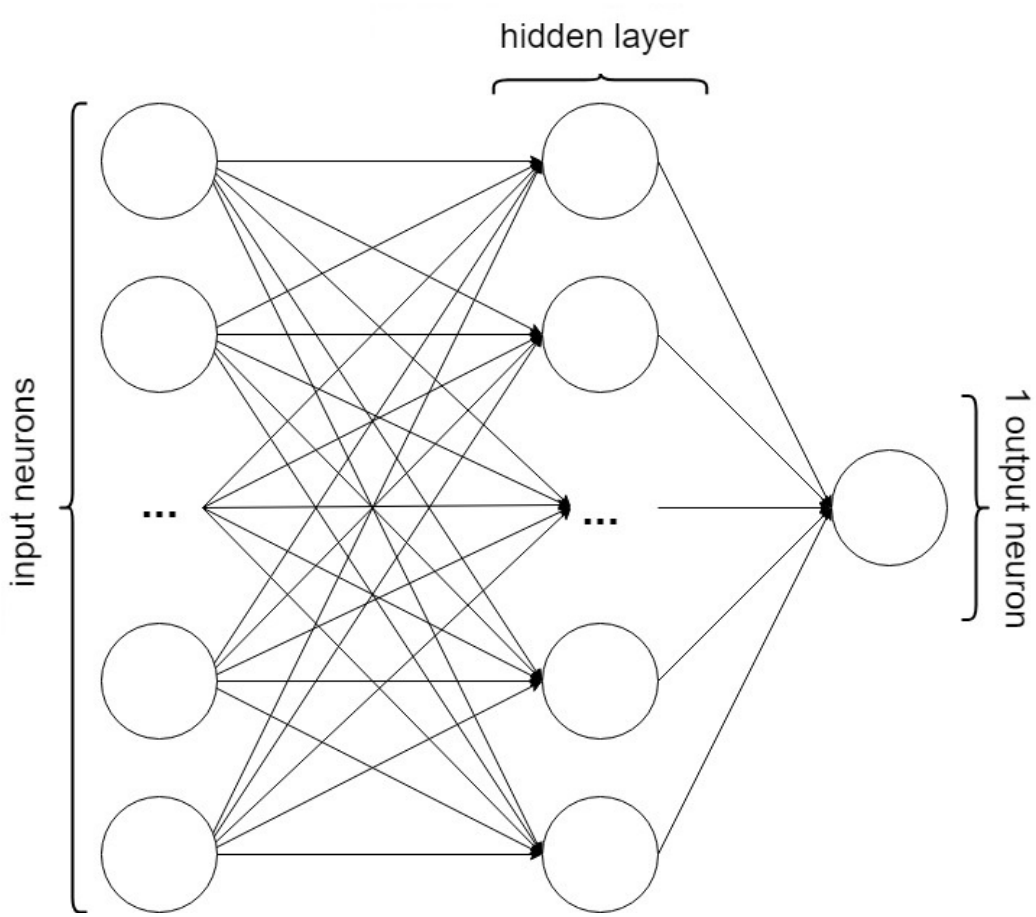

Figure 2: *Standard* network.

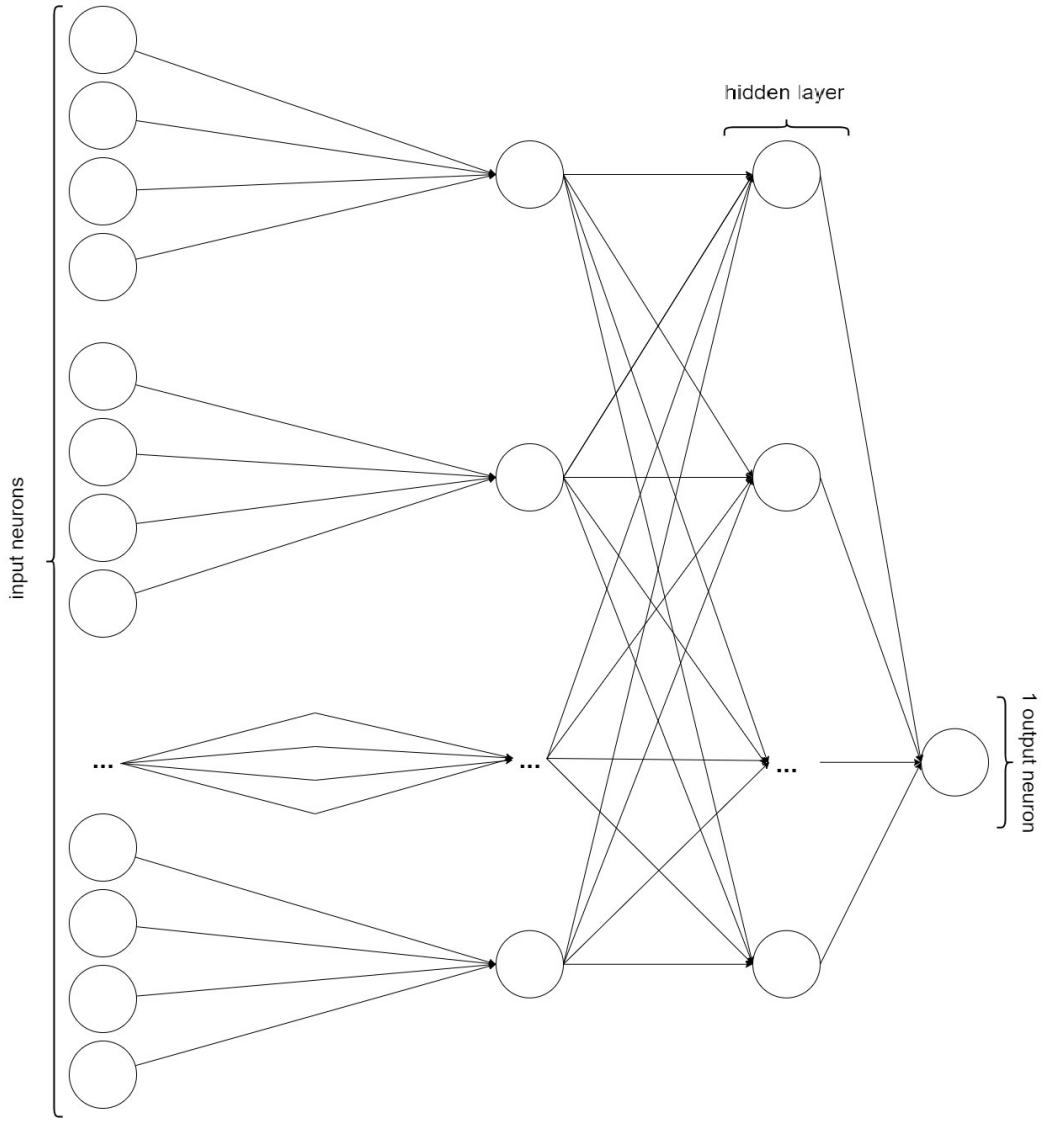

Figure 3: *Ad-hoc* network: input neurons are grouped by four, indicating the observations over a year of the same variable.

Algorithm 1: Adaptive hidden neurons computation for Neural Networks
Initialization: observational data set
The optimal network topology w.r.t. the error defined in Eq. 2. b in $1, \ldots, \#$ sub-sampling runs $X_b \leftarrow$ Dataset at b^{th} sub-sample $Xtrain_b \leftarrow$ Training set at b^{th} sub-sample run $Xtest_b \leftarrow$ Test set at b^{th} sub-sample run i in $1, \ldots, \#$ of hidden layer $k \leftarrow 0$ $j \leftarrow 1$ $k \leq K$ **train** net Net_{ij} on $Xtrain_b$ **compute** $RMSE_{ij}$ error with Eq. 2 on the $Xtest_b$ $Overall_{ij} \leq$ **all** $RSME_{ij}$ $BestNet \leftarrow Net_{ij}$ $Neurons \leftarrow (i, j)$ $Overall_{ij} > RSME_{Neurons}$ $k{+}{+}$ $k \leftarrow 0$ $j{+}{+}$ $BestNet_b \leftarrow BestNet$ $Neurons_b \leftarrow Neurons$ **return** BestNet, RSME=Eq. 2, Neurons

5.1.2 Training and test set

In our experiments we are exploiting the *supervised* learning, meaning that during the *learning* phase, for each input pattern, we are also providing the desired output value, that in our scenario corresponds to CAR: all other indicators considered in Table 3 will define the input pattern for each financial institution. In order to grasp the time dynamics, for each input indicator i we are providing to the network the value of the indicators collected at time t, together with the values collected at time $(t-1)$, $(t-2)$, and $(t-3)$. During the Neural Network learning we have to identify two disjoint sets of observations out of the overall 34944 observations: the *training set*, that will be used to determine the synapses' weights, and the *test set*, that will be used to determine the network performance and to stop the learning. According to [19], we have decided to split the overall data by randomly allocating the 70% of its observations to the *training set*, and the remaining 30% to the *test set*. This random allocation has been repeated 50 times, each time determining a different *train-test* partition. We have then run our Neural network approaches on all obtained partitions, and in what follows we are reporting, for each Neural approach, the average and standard deviation statistics over the 50 partitions.

5.2 Linear Regression

Linear Regression is used to model the relationship between two or multiple parameters by fitting a linear equation on the observed data. Usually, this is done using the *least-square* regression that minimizes the sum of squares of the vertical deviation from each data point on the line. The algorithm aims to reduce this sum by selecting the most appropriate constants in the equation representative of the regression line.
A Linear Regression line has an equation of the form $Y = a + bX$ where X is the explanatory variable and Y is the dependent variable. The slope of the line is b, and a is the intercept (the value of y when $x = 0$) [60].

5.3 Generalised Linear Models (GLMs)

The basic Linear Regression predicts a certain value as a linear combination of a specific set of observed values, meaning that a change in one or multiple predictors affects the response variable. However, for complex data, Linear Regression is not very effective, and in these cases one may resort to Generalized Linear Models [41], that allow response variables to have arbitrary distributions (rather than normal distributions), and define an arbitrary function of the response variable (the link function) to vary linearly with the predictors, rather than assuming that the response itself must vary linearly. A generalized Linear Model (GLM) consists of three elements: Linear predictor; Link function; Probability distribution or exponential family. The linear predictor is the linear combination of parameter b and explanatory variable x. The link function is what links the linear predicted and the probability distribution: there are many link functions and usually, they are used depending on the features of data we are trying to predict and in which range are we expecting it to be.

5.4 Evaluation of the model

Once we build a model through Linear Regressions (or GLMs), we need to measure the correctness of this model. This can be done via different statistical measures, that are used to benchmark the performance of predictive models.

5.4.1 Root mean squared error

The root mean squared errors represents the standard deviation of the prediction errors. by that, we mean that it tells us how concentrated the data is around the line that we predict. To calculate it we can take the square root of the Mean squared errors.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y_i})^2}.$$

where n is the number of predictions, Y_i the observed values, and $\hat{Y_i}$ being the actual prediction of that variable.

5.4.2 R-Squared

The R-Squared represents how well the data fits on the regression line. More

generally, it is used to analyze how the difference in one variable can be explained by other variables. In the case of regression, we can reason in percentages and say that the closer the measure is to 1 the closer the points are to the regression line up to 1 where 100% of the points are on the line. Generally, this would mean that the higher R-squared is the better the results we have but this can be false in some edge cases. It is calculated by squaring the correlation coefficient calculated with this formula

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

where SS_{res} is the sum of squares of residuals, and SS_{tot} is the residual sum of squares.

5.4.3 F-statistic

The *F-statistic* in a regression is a value that represents how well you improved the regression line compared to a regression line with all the coefficients $= 0$. if your model significantly improved the model fit then you will get a better *F-statistic*. But before taking into account the *F-value* one must first look at the *P-value* that is calculated at the same time as the *F-statistic*. With the *F-statistic* calculation comes the *P-value*. Usually the *P-value* is looked at before taking the *F-statistic* into account. If the *P-value* is lower than the alpha level, then we can reject the null hypothesis and we can consider the *F-value*, otherwise the *F-values* is worthless.

6 Results and discussion

In this section we report the principal results to build a model that is performs well on our benchmarks.

All Neural approaches have been implemented in Python, exploiting the library Tensor-Flow. Experiments have been run on a on a cluster with AMD Opteron 2216 dual core CPUs running at 2.4 GHz with 2x1 MB L2 cache and 4 GB of RAM under Cluster Rocks distribution built on top of CentOS 5.3 Linux.

Table 4 reports the RMSE of the experiments run with both the *Standard* and *Ad-hoc* networks. We have performed 50 runs of the adaptive procedure devised in Algorithm 2 and reported the minimum, maximum, mean, median, and standard deviation of the RMSE distribution, for each possible instantiation of the pair $[network, model]$.

	Standard Reduced model	Standard Full model	Ad-hoc Reduced model	Ad-hoc Full model
Min	0.070	0.071	0.040	0.040
Max	0.074	0.076	0.045	0.051
Mean	0.072	0.073	0.042	0.045
Std	0.001	0.001	0.001	0.003
Median	0.072	0.073	0.042	0.046

Table 4: RMSE of the experiments run with the *Standard* and *Ad-hoc* networks. For each columns. For each column, statistics over 50 runs of the proposed adaptive procedure are reported.

As a first remark, we can see that the pre-processing operation have a valid role in improving the networks' performances, in both $Standard$ and *Ad-hoc* networks. Then, we can see that the *Ad-hoc* networks' error is lower than the $Standard$ one. This confirms the results found by [17] and [21], in which authors exploit the fact that the *Ad-hoc* network is able to grasp the temporal dependence of inputs.

Then, we are presenting the results obtained with Linear Regression and with GLMs (both developped using the Sci-kit learn library on Python[51]), and then, we are comparing them with the Neural Network approach devised in section 5.1. As a first experiment, we have implemented a regression approach over the whole dataset, and the results are shown in table 5: in this table we report the R^2 relative to experiments performed with Linear Regression and GLMs, along with two variants: the Adjusted R-Squared (that takes into account the number of predictors) and the Predicted R-Squared (that takes into account overfitting). Please notice that for GLM we have used the *pseudo R-squared* (for sake of definition) and that we also report the *P-value* of the *F-test* (i.e., the probability of obtaining an *F-statistic* value that is greater than the model's *F-value*, under the null hypothesis that the regression model is not significant: low positive values identifies good fit). We have performed experiments by using as predictors both the total set of variables identified after the pre-processing phase (identified by the entry *Whole set of data* in the table), and a limited set of predictors composed of all predictors that are significant for the regression according to their *P-value* (identified by the entry *All observations, limited set of predictors*). We see that the reduction of predictors does not improve the goodness of the fit according to the different metrics used, so in what follows we are using the whole set of predictors. In this direction, we remark that Generalised linear models lead to a better $R-squared$, but this comes at the cost of a higher overfitting, as witnessed by the lower value of the $PredictedR-squared$. Linear Regression instead, leads to a worse (but still acceptable) $R-squared$, but its difference with the $PredictedR-squared$

is lower, showing a better robustness of the approach. Please notice that all regressions are significant according to the *P-value* of the *F-test*. We recall that these experiments have been performed on the whole set of data. In what follows we will describe experiments performed on different partitions of $training/test$ sets, in order to compare these approaches with our Neural Network approach.

Table 5: Linear Regression and GLMs over the whole set of data: measures of the goodness of the fit.

Data Used	Model	statistic	values
All observations, limited set of predictors	Linear Regression	Predicted R-squared	0.77
		R-squared	0.78
		Adj R-squared	0.78
Whole set of data	Linear Regression	Predicted R-squared	0.77
		R-squared	0.78
		Adj R-squared	0.78
		P-value of the F-test	0.0
Whole set of data	GLM	Predicted R-squared	0.71
		Pseudo R-Squared	0.91
		P-value of the F-test	0.0

Please notice that the previous results have been obtained on the whole set of data. In order to test the generalization capability of our approach, we have split the whole set of data in 50 different $training/testing$ partitions (i.e., the same partitioning generated in Section 5.1.2), built our model to fit the training set, and assessed the goodness of the fit on the test set. This has been done for all regression detailed in Table 5, along with the Neural approaches devised in Table 4, and the goodness of fit (assessed by the $R - squared$) has been reported in Table 6, as computed on the error distributions displayed in Table 4.

By looking the results, we see that the best results are offered by the *Ad-hoc* Neural Networks, but also the Standard network performs fairly well. This is due to the generalisation skill of the network, able to prevent the overfitting we have found on the aforementioned regression approaches. This confirms the goodness of our adaptive procedure, that has been proposed by [17] in a different context, but that can be tailored to the different application scenarios.

Table 6: Experiments with Linear Regression, Generalised Linear Models, and Neural Networks (Standard and Ad-hoc). Statistics of the R-Squared computed on the test sets of 50 different $training/testing$ partitions.

Model	statistic	values
Linear Regression	min	0.76
	max	0.80
	mean	0.77
	stdd	¡ 0.01
GLM	min	0.69
	max	0.72
	mean	0.70
	stdd	¡ 0.01
Standard NN	min	0.71
	max	0.90
	mean	0.81
	stdd	0.15
Ad-hoc NN	min	0.64
	max	0.92
	mean	0.84
	stdd	0.13

7 Concluding remarks

The bankruptcy of banks may lead to a huge catastrophic effect over the overall economy, since the contagion effect it may trigger could lead to a generalised overall crisis. To this extent, the activity of banking supervision, and the role and the authority of bank regulation, play a big role, since they may prevent (or reduce the effect of) the banks' bankruptcy. Although aimed to different targets, many of these supervision exercise are designed to maintain a sufficient level of capital adequacy: in a way, the banks have to allocate specific reserves to face expected losses and to protect themselves from excessive credit expansion. Bank regulations impose constraints over these reserves, even though banks operate themselves preventively against unexpected crises, and their reserves are often higher than the ones imposed by the regulations. The CAR is one of the main indicators monitored by the banks themselves and by the supervising authorities in order to assess the bank health, and in our contribution we have devised an adaptive Neural Network approach to predict the CAR, and compared the obtained results with standard approaches such as *Linear Regression* and *Generalised Linear Models*.

Results show that Neural Networks may be successfully used to predict the CAR, and that their outcomes compare favourably with standard methods when used jointly with meaningful pre-processing operations. In future research, modern recurrent neural networks (Long Short-Term Memory or Gate Recurrent Units) and 1D-convolutional Neural Networks could be used to exploit the time dependency of the data.

References

[1] S. Park S. Peristiani A., Estrella. Capital ratios as predictors of bank failure. *Economic Policy Review*, 6:33–52, 02 2000.

[2] V. Siakoulis E. Stavroulakis N. E. Vlachogiannakis A., Petropoulos. Predicting bank insolvencies using machine learning techniques. *International Journal of Forecasting*, 36(3):1092 – 1113, 2020.

[3] V. V. Acharya, D. Pierret, and S. Steffen. Introducing the "leverage ratio" in assessing the capital adequacy of european banks. *ZEW Discussion Und Working Paper*, 49(621):460–482, 2016.

[4] T. Adrian, J. Morsink, and L. B. Schumacher. Stress testing at the imf. Technical report, International Monetary Fund, 2020.

[5] Z. Affes and R. Hentati-Kaffel. Forecast bankruptcy using a blend of clustering and mars model: Case of us banks. *Annals of Operations Research*, 281(1):27–64, 2019.

[6] N. M. Al-Sabbagh. *Determinants of capital adequacy ratio in Jordanian banks*. PhD thesis, Yarmouk University, 2004.

[7] A. Alfadli and H. Rjoub. The impacts of bank-specific, industry-specific and macroeconomic variables on commercial bank financial performance: evidence from the gulf cooperation council countries. *Applied Economics Letters*, 27(15):1284–1288, 2020.

[8] F. Audrino, A. Kostrov, and J.P. Ortega. Predicting u.s. bank failures with midas logit models. *Journal of Financial and Quantitative Analysis*, 54(6):2575–2603, 2019.

[9] A. Bahrammirzaee. A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8):1165–1195, 2010.

[10] H. B. Barlow. Unsupervised learning: introduction. In G. E. Hinton and T. J. Sejnowski, editors, *Unsupervised Learning: Foundations of Neural Computation*, pages 1–17. Bradford Company Scituate, MA, USA, 1999.

[11] J. R. Barth and G. Caprio. Approaches to bank supervision. *Political institutions and financial development*, page 156, 2008.

[12] L. Bateni, H. Vakilifard, and F. Asghari. The influential factors on capital adequacy ratio in iranian banks. *International Journal of Economics and Finance*, 6(11):108–116, 2014.

[13] C. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 2005.

[14] K. Bourkhis and M. S. Nabi. Islamic and conventional banks' soundness during the 2007–2008 financial crisis. *Review of Financial Economics*, 22(2):68 – 77, 2013.

[15] J. A. Chattha and S. Archer. Solvency stress testing of islamic commercial banks: assessing the stability and resilience. *Journal of Islamic Accounting and Business Research*, 2016.

[16] M. Čihák. Introduction to applied stress testing. *IMF Working Papers*, pages 1–74, 2007.

[17] M. Corazza, D. De March, and G. di Tollo. Design of adaptive elman networks for credit risk assessment. *Quantitative Finance*, 21(2):323–340, 2021.

[18] R. Dakovic, C. Czado, and D. Berg. Bankruptcy prediction in norway: a comparison study. *Applied Economics Letters*, 17(17):1739–1746, 2010.

[19] G. di Tollo. Reti neurali e rischio di credito: stato dell'arte e analisi sperimentale. Technical Report R-2005-003, Dipartimento di Scienze, Università "G. D'Annunzio" Chieti–Pescara, 2005.

[20] P. Dua and H. Kapur. Macro stress testing and resilience assessment of indian banking. *Journal of Policy Modeling*, 40(2):452–475, 2018.

[21] E.Angelini, G.di Tollo, and A. Roli. A neural net approach for credit-scoring. *Quarterly Review of Economics and Finance*, 48:733–755, 2008.

[22] LM. Fu. *Neural Networks in Computer Intelligence*. McGraw-Hill, Inc., USA, 1994.

[23] P. Gai, A. Haldane, and S. Kapadia. Complexity, concentration and contagion. *Journal of Monetary Economics*, 58(5):453–470, 2011.

[24] N. Gambetta, M. A. García-Benau, and A. Zorio-Grima. Stress test impact and bank risk profile: Evidence from macro stress testing in europe. *International Review of Economics & Finance*, 61:347–354, 2019.

[25] M. G. Gulaliyev, N. P. Ashurbayli-Huseynova, A. A. Gubadova, B. N. Ahmedov, G. M. Mammadova, and R. T. Jafarova. Stability of the banking sector: deriving stability indicators and stress-testing. *Polish Journal of Management Studies*, 19, 2019.

[26] L. H. Wang J. P. Yin Chen P. H. Chen H. F. Zhang H. F., Zhang. Performance of the levenberg–marquardt neural network approach in nuclear mass prediction. *Journal of Physics G: Nuclear and Particle Physics*, 44(4):045110, mar 2017.

[27] A. Hadjixenophontos and C. Christodoulou-Volos. Financial crisis and capital adequacy ratio: A case study for cypriot commercial banks. *Journal of Applied Finance and Banking*, 8(3):87–109, 2018.

[28] A. Haldane. Constraining discretion in bank regulation. *Central Banking at a Crossroads*, page 15, 2013.

[29] M. K. Hassan, O. Unsal, and H. E. Tamer. Risk management and capital adequacy in turkish participation and conventional banks: A comparative stress testing analysis. *Borsa Istanbul Review*, 16(2):72–81, 2016.

[30] JB Heaton, N. G. Polson, and J. H. Witte. Deep learning in finance. *arXiv preprint arXiv:1602.06561*, 2016.

[31] H. Husna and R. Rahman. Financial distress–detection model for islamic banks. *International Journal of Trade, Economics and Finance*, pages 158–163, 01 2012.

[32] A. Jamali. Modeling effects of banking regulations and supervisory practices on capital adequacy state transition in developing countries. *Journal of Financial Regulation and Compliance*, 2019.

[33] K. Kumar A. Gepp K., Halteh. Financial-distress prediction of islamic banks using tree-based stochastic techniques. *Managerial Finance*, Special Issue in the Role of Islamic Finance in Mainstream Finance, 08 2017.

[34] R. A. A. Karim. The impact of the basle capital adequacy ratio regulation on the financial and marketing strategies of islamic banks. *International Journal of Bank Marketing*, 1996.

[35] A. Khashman. Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Systems with Applications*, 37(9):6233 – 6239, 2010.

[36] S. B. Kotsiantis. Supervised machine learning: A review of classification techniques. *Informatica*, 31(3):249–268, 2007.

[37] K. Kočišová and M. Mišanková. Discriminant analysis as a tool for forecasting company's financial health. *Procedia - Social and Behavioral Sciences*, 110:1148 – 1157, 2014. The 2-dn International Scientific conference Contemporary Issues in Business, Management and Education 2013".

[38] N. Laila and F. Widihadnanto. Financial distress prediction using bankometer model on islamic and conventional banks: Evidence from indonesia. *International Journal of Economics and Management*, 11:169–181, 01 2017.

[39] D. Martin. Early warning of bank failure : A logit regression approach. *Journal of Banking & Finance*, 1(3):249–276, November 1977.

[40] D. Mayes and H. Stremmel. The effectiveness of capital adequacy measures in predicting bank distress. *SUERF*, 2014/1, 02 2014.

[41] P. McCullagh and J.A. Nelder. *Generalized Linear Models, Second Edition*. Chapman and Hall/CRC Monographs on Statistics and Applied Probability Series. Chapman & Hall, 1989.

[42] M. Mehreen, Maran M., S. Ariffin A. Karim, and Amin J. Proposing a multidimensional bankruptcy prediction model: An approach for sustainable islamic banking. *Sustainability*, 12:3226, 04 2020.

[43] E. Montero, M. C. Riff, and B. Neveu. A beginner's guide to tuning methods. *Appl. Soft Comput.*, 17:39–51, April 2014.

[44] G. E. Morgan. On the adequacy of bank capital regulation. *Journal of Financial and Quantitative Analysis*, 19(2):141–162, 1984.

[45] D. M. Nachane and S. Ghosh. Credit rating and bank behaviour in india: Possible implications of the new basel accord. *The Singapore Economic Review*, 49(01):37–54, 2004.

[46] A. K. NOVOKMET and A. BANOVIĆ. Why do the minimum capital adequacy ratios vary across europe? *Journal of Applied Economic Sciences*, 11(3):41, 2016.

[47] H. Oloo, M. Wanjiru, and K. Newell-Jones. Female genital mutilation practices in kenya: the role of alternative rites of passage. a case study of kisii and kuria districts. 2011.

[48] T. Loughran B. McDonald P., Gandhi. Using annual report sentiment as a proxy for financial distress in u.s. banks. *Journal of Behavioral Finance*, 20(4):424–436, 2019.

[49] J. Park, M. Shin, and W. Heo. Estimating the bis capital adequacy ratio for korean banks using machine learning: Predicting by variable selection using random forest algorithms. *Risks*, 9(2), 2021.

[50] J. Park, M. Shin, and W. Heo. Estimating the bis capital adequacy ratio for korean banks using machine learning: Predicting by variable selection using random forest algorithms. *Risks*, 9(2):32, 2021.

[51] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[52] A. Petropoulos, V. Siakoulis, K. Panousis, T. Christophides, and year = 2020 month = 09 pages = title = A Deep Learning Approach for Dynamic Balance Sheet Stress Testing Chatzis, S.

[53] V. Ravi and C. Pramodh. Threshold accepting trained principal component neural network and feature subset selection: Application to bankruptcy prediction in banks. *Applied Soft Computing*, 8(4):1539 – 1548, 2008. Soft Computing for Dynamic Data Mining.

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

[55] B. Robitaille, B. Marcos, M. Veillette, and G. Payre. Modified quasi-newton methods for training neural networks. *Computers Chemical Engineering*, 20(9):1133–1140, 1996.

[56] R. Rojas. The backpropagation algorithm. In *Neural networks*, pages 149–182. Springer, 1996.

[57] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.

[58] N. Vunjak, N. Milenković, J. Andrašić, and M. Pjanić. Stress test model for measuring the effects of the economic crisis on the capital adequacy ratio.

[59] D. Worrell. Stressing to breaking point: Interpreting stress test results. 2008.

[60] X. Yan and X. G. Su. *Linear Regression Analysis: Theory and Computing*. World Scientific Publishing Co., Inc., USA, 2009.