Turkey Earthquake Prediction with Deep Learning Algorithms

MSc Research Project
MSc in Data Analytics

Omer KAS
Student ID: x20244100

School of Computing
National College of Ireland

Supervisor: Abdul SHAHID

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Omer KAS
Student ID:	x20244100
Programme:	MSc in Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Abdul SHAHID
Submission Due Date:	08/14/2023
Project Title:	Turkey Earthquake Prediction with Deep Learning Algorithms
Word Count:	XXX
Page Count:	25

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	14th August 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Turkey Earthquake Prediction with Deep Learning Algorithms

Omer KAS
x20244100:

Abstract

Earthquakes, a natural phenomenon prevalent in seismically active regions, pose a grave threat to human lives and infrastructure. Despite today's technological advancements, a definitive earthquake prediction method remains elusive. However, researchers across diverse scientific fields are diligently studying past earthquake records in hopes of uncovering discernible patterns. To anticipate impending earthquakes, comprehensive investigations are conducted, drawing expertise from various disciplines. Notably, the rise of information technologies has steered attention towards deep learning, a subset of artificial intelligence, as a means to achieve accurate predictions in this complex process. In this research study, research is carried out on a possible future earthquake prediction model using deep learning architectures with the data of earthquakes that have occurred in Turkey. The topic of study includes a model investigated by using information such as time, latitude, longitude, magnitude and depth of the catalog data of earthquakes that have occurred in Turkey and deep learning algorithms. Long-Short Term Memory (LSTM) architecture, which is one of the Recurrent Neural Network processes, is used to predict the time of occurrence of the earthquake that may occur in the research. In the process of developing the model, the RNN model demonstrates superior prediction accuracy compared to the LSTM model. The RNN model achieves lower values for all evaluation metrics: MSE (95.13 vs. 195.54), RMSE (9.75 vs. 13.98), and MAE (4.70 vs. 5.71). This underscores the RNN model's better performance, reflected by consistently reduced error metrics across all measures.

Keywords: Earthquake, Deep Learning, LSTM, Turkey

1 Introduction

1.1 Backround and Motivation

An earthquake is an event where the earth's crust fractures suddenly, causing vibrations that propagate as waves through the environment, resulting in shaking. Earthquakes, earthquakes occur in seismically active regions since the formation of the world, resulting in the death of millions of people and the destruction of their habitats. In earthquakes, rain, wind, and similar natural events are accepted. It is possible to predict rain, wind, and similar natural events with today's technologies and to minimize the damages by taking various precautions. It is a very difficult issue to take precautions because the time and magnitude of the earthquake is a destructive natural event that cannot be predicted. Estimating the occurrence time of the earthquake is an extremely important issue in order to overcome the material and moral damages with the least damage during the earthquake process.

Mohankumar and Sangeetha (2018) Prediction of earthquakes greatly benefits from the research and analysis of earthquakes. Pandit and Biswal (2019)However, accurate estimation of the time, magnitude, and location of earthquakes is a difficult issue today. Earthquake forecasting is the idea of creating a warning mechanism that aims to detect earthquakes in advance and take measures to reduce the risk of death and destruction. Achieving good results in earthquake prediction depends on a good analysis of previous earthquakes.

Zakeri and Pashazadeh (2015) The number of forecasting studies using smart calculation methods for earthquake prediction is increasing. Although there are various studies on earthquake prediction among different scientific disciplines, there is no study that gives a definite result today. In addition, the most popular studies in the field of earthquake forecast are focused on earthquake forecast studies with deep learning architectures that emerged with the development of statistics and artificial intelligence technologies. DeVries et al. (2018) Deep learning in particular, as well as other applications of artificial intelligence, have gained mainstream acceptance, raising new prospects for earthquake prediction and resolution.

Sen (2023)The earthquake has an impact on numerous nations or areas. The areas that will be most severely impacted are those that have traveled through countries or regions where earthquake faults have occurred. The fact that Turkey is in an earthquake zone is something that must be acknowledged. These places are even more well-known because several parts of Turkey are situated in an earthquake zone. Turkey's geographic location results in a variety of fault lines. Sen (2023) Major fault lines like the North Anatolian Fault Line, East Anatolian Fault Line, West Anatolian Fault Line, and South Anatolian Fault Line are present in Turkey. These fault lines pose a significant earthquake risk across most of the country.

Türkiye is constantly exposed to earthquakes due to the earthquake zone it is located in. In this regard, the last Kahramanmaraş earthquake, which occurred in Turkey and is called "dual earthquakes", which is rare in the world, took place in our country on 06 February 2023 "04.17". More than 50,000 people lost their lives in these earthquakes. Sen (2023) The magnitude of the damage inflicted by the earthquake may be in excess of 100 billion dollars, according to statements made by Louisa Vinton, the United Nations Development Programme's Turkey Regional Representative. This is equivalent to 11.6% of Turkey's projected 2021 Gross domestic product. The impact of an earthquake on Turkey would be quite terrible, as can only be seen from these numerical facts.

1.2 Research Questions

Ergunay (2007) Turkey is situated within the Mediterranean-Alpine-Himalayan seismic belt, which is recognized as one of the world's most active seismic zones. Altun (2018) Existing earthquake zones on 93% of Turkiye's land. Knowing the whereabouts and duration of occurrence of an earthquake in a region located in an earthquake zone at this level is an extremely important issue.

In this research study, a full-time earthquake prediction model can be developed with the information such as time, latitude, longitude, magnitude and depth of the catalog data of earthquakes occurring in Turkey. In the research, Long-Shor Term Memory (LSTM) model, one of the recursive neural network processes, will be used to predict the occurrence time of earthquakes. The results of this model will be compared with the results of other deep learning models such as Recurrent neural network (RNN), Graded Recurrent Units (GRU), Bidirectional LSTM (BiLSTM).

1.3 Research Objectives

The objective is to create a model for predicting earthquakes based on historical information on earthquakes that happened in Turkey between 1900 and 2018.

It is aimed to develop a model on earthquake occurrence time by using the Long-Short Term Model (LSTM) neural networks approach, the relationship between earthquake data (date, time, latitude, longitude, magnitude, depth) and earthquake occurrence time.

The developed Long-Short Term Model (LSTM) neural network approach model results will be compared with the results of Bidirectional LSTM (BiLSTM), Gated Recurrent Units (GRU), Recurrent neural network (RNN) and it is aimed to reveal which neural network approach will result in a more successful prediction on the data.

2 Related Work

2.1 Earthquake

Earthquake is a natural event that underscores the unpredictability of the supposedly stable ground beneath our feet. It serves as a reminder that even the ground we consider solid can shift suddenly, causing widespread destruction and posing threats to human life and structures. With today's technologies, many natural events can be predicted. However, an adequate technology for earthquake forecast has not been developed yet. Since the subject of earthquake prediction is a very complex and specific subject, studies in this field do not give very precise results. However, each study serves as a source for the next study. However, it is thought that with the rapid progress of today's technologies, studies that will give very good results in earthquake prediction will be carried out in the future. Here, many studies on earthquake prediction are given.

Isci (2008)The phenomenon of shaking the environments through which these vibrations, which are caused by the energy that emerges suddenly due to these fractures in the Earth's crust, spread as waves and pass through, is called an earthquake. The lithosphere, which surrounds our world from the outside, has a structure that is divided into many small particles called plates or plates. Ketin (1994) Movements between these plates that make up the lithosphere have provided the formation of mountains and continents. It is known that movements between plates take many years. The rapid movements of these plates are called earthquakes.

The ruptures in the region of the plate where the shaking event occurs are called faults. It is known that while fault lines can sometimes be detected with the eye on the ground, sometimes these lines cannot be detected with the eye below the ground surface. However, even if the fault lines are not a definite point direction with today's technology, it is possible to detect them regionally.Karakus (2014) This gives us the opportunity to know where the earthquake may occur but does not give us the opportunity to predict when it will happen. Buyukkaracigan (2016)The majority of earthquakes in the world happen on slender bands at plate borders when these plates push against one another. Isci (2008)If we look at the current Earthquake Risk Maps and Active Fault Lines Maps of

Turkey, 96 6% of the total surface area in Turkey is I., II., III. and IV. degree seismic zones. 986%the total population in Turkey, that is, almost all of them, live under earthquake risk even though they are under different earthquake zones.

Isci (2008)A measure of the energy released at the source of a magnitude earthquake; magnitude is a measure of the effects of an earthquake on structures and people. The Magnitude of the earthquake is defined as the logarithm of the amplitude of the earthquake waves on the seismogram recorded in a certain time period.Utkucu et al. (2023) It is defined as the measure of the effect of an earthquake that occurs anywhere on the earth at a point where it is felt. While the intensity does not give accurate information about the magnitude of the earthquake at its source, it reflects the damage caused by the earthquake depending on the factors mentioned above.Bikce (2017)Turkey is one of the three most significant earthquake zones in the world and is situated on the Alpine-Himalayan earthquake fault zone in the Mediterranean region.

2.2 Earthquake Prediction Deep Learning

Since disasters such as earthquakes do not give warning beforehand, it is a really difficult issue to prevent losses at the time of their occurrence. Although earthquake prediction is a difficult subject, it is a slightly more interesting subject compared to other types of disasters. With the rapid development of information technologies and advances in the field of machine learning, artificial intelligence-based studies show efficiency in earthquake prediction. In the literature survey, there are many studies on earthquake prediction based on artificial neural networks. It is thought that with the continuity of studies on earthquake prediction, a certainty in earthquake prediction will be reached. Although studies on earthquake prediction were limited in the past, research on earthquake prediction is increasing today.

Holliday et al. (2007)As an alternative method for earthquake prediction, it is suggested to consider the past occurrence rates of small earthquakes and regions of high seismic activity or stagnation. These regions are considered places where large earthquakes are expected. Wiemer and Wyss (1994)Another important statistic on earthquake behavior is that the leading seismic stagnation assumption can be defined as a significant reduction in average seismicity compared to the activity rate of current earthquakes in an earthquake zone.

Holliday et al. (2007)Earthquake forecasting methods can generally be categorized into two classes, the first taking into account experimental measurements of precursor changes, while the second approach is to use the statistical generation of seismic activity. Bodri (2001)demonstrated that the forecast of large earthquakes can be significantly influenced by changes in seismic activity patterns. Ozturk and Alkan (2022)One of the most preferred parameters in earthquake prediction studies is the Gutenberg-Richter (G-R) relationship, which reveals the magnitude-frequency relationship of earthquake distributions.

Sebatli-Saglam and Cavdur (2022)Optimal outcomes concerning performance metrics were achieved when applying Bayesian regularization and the Levenberg-Marquardt training algorithm for forecasting earthquake magnitude. This was contingent on factors such as the quake's depth and magnitude, as well as the distance of affected individuals from the epicenter. The utilized approach involved a feed-forward back-propagation artificial neural network.

Alarifi et al. (2012)The nonlinear relationships of a feedforward neural network are more predictive than statistical methods.Reyes et al. (2013)The proposes the application of the Hurst exponent to predict the predictability of the earthquake series.Moustra et al. (2011)The research findings on estimating the magnitude of impending seismic activity following a delay between seismic electrical signals and seismic activity emphasise the finding that neural networks, when trained using relevant data, can generalise and predict unknown seismic events reasonably accurately.

Narayanakumar and Raja (2016)Using a supervised learning algorithm for Multilayer Perceptron training, it is stated that when the values obtained are compared with the earthquake results, it is more successful in predicting earthquakes between 4.0-6.0 compared to other models. Mallouhy et al. (2019)An earthquake dataset was subjected to eight different algorithms, and the results show that KNN, Random Forest, and MLP are the best by creating the fewest false outputs (FP) and properly classifying a greater proportion of outputs than SVM, KNN, and MLP.

Niksarlioglu and Kulahci (2013)Artificial neural networks provide a number of advantages over numerical methods in unstructured systems, including the potential for quick optimisation and the lack of a need for mathematical formulas to connect the data in the input-output processes. However, neural networks also have the drawback of requiring a lot of data to get accurate findings. Yildirim (2010)In the research on the clustering of earthquake data that happened throughout Turkey, it is claimed that the KNN method performs better than the conventional k-Means algorithm.Berhich et al. (2022)Allows clustering to focus its models independently on each region.

Sheng et al. (2015)The accuracy of the estimates may be increased by breaking the research zone into smaller sub-regions based on sufficient seismicity data. Bilen et al. (2015)It is emphasized that the occurrence of the earthquake is predicted by using the seismic data taken from the Polish coal mines, and the use of the KNN algorithm in earthquake prediction will give successful and fast results.

Jang (1993)The Adaptive Neuro-Fuzzy Inference System (ANFIS) stands as one of the initial hybrid neuro-fuzzy systems for function prediction. The outcomes of the correlation factor indicate that ANFIS exhibits a strong capacity for earthquake prediction. Kamath and Kamat (2017)The ANFIS model, integrating artificial neural network (ANN) and fuzzy inference systems (FIS), excels in its capability to transform numerical forecasts into linguistic terms through the utilization of ANN and fuzzy grading mechanisms.

Mirrashid et al. (2016)There is a notion that the Neuro-Fuzzy Inference System (ANFIS) holds promise as a valuable tool for predicting the seismic moment of upcoming earthquakes.Shibli (2011)According to the principle of conservation, the cumulative moment curve of earthquakes in the northern hemisphere is thought to equate the cumulative moment curve in the southern hemisphere of the planet.

Asencio-Cortes et al. (2015)Classification trees and random forest methods have been successful in principal component analysis (PCA), where the usage of principle component analysis has been recommended to reduce data dimensionality and to build new datasets.Cilli (2007)The use of principal component analysis (PCA), which tries to retain the changes of the associated attributes in the samples, underlines that it is a transformation approach to minimise the number of attributes by transferring the attributes to a different dimension.

Asim et al. (2020)In the earthquake research in Cyprus, Random Forests made the most accurate predictions at magnitudes 3.0 and 3.5, while Support Vector Machines gave better results than other methods at magnitudes 4.0 and 4.5. It is seen that these methods can be effective in short-term earthquake predictions.Vasti and Dev (2020) study involving data categorization through diverse machine learning techniques including K-nearest neighbor, support vector machine, Naïve Bayes, and random forest, it was observed that the random forest algorithm outperformed the rest in terms of accuracy

Jozinovic et al. (2020)The findings indicate that the strength of ground vibrations can be predicted utilizing a convolutional neural network by directly processing raw waveform data, bypassing the necessity for earthquake source details like location and magnitude. This CNN-based approach results in a reduced bias within the CNN model as compared to alternative techniques.Wang et al. (2009) The utilization of the Radial Basis Function (RBF) artificial neural network in earthquake prediction has demonstrated favorable outcomes. The RBF artificial neural network model displayed effective performance within the domain of earthquake prediction.

Azam et al. (2014)Research on covering different AI-related strategies and cross-checking their reliability also suggests that a single AI analysis application would be insufficient to predict outcomes at a good level.Iatan (2015)In the research on better understanding the Probabilistic Neural Network (PNN) to be applied to solve the problem of earthquake magnitude estimation, it is said that, unlike neural networks, no learning rule is required for a PNN.

2.3 Earthquake Prediction With LSTM

LSTM is an RNN suggested by Hochreiter and Schmidhuber (1997).Recurrent Neural Networks (RNNs) are a type of artificial neural network designed with memory capabilities. RNNs are suitable for analyzing time series data due to their ability to capture interdependencies within the data. However, conventional RNNs have limitations. They encounter issues like vanishing and exploding gradients, rendering them difficult to train effectively. Developed by Hochreiter and Schmidhuber (1997), LSTM is designed to solve some of the difficulties in traditional recurrent neural networks (RNN).

Gavcar and Metin (2021)The LSTM memory cell consists of weight parameters for input, output, and input time steps. Input weights are utilized to scale the input at the current time, while output weights are applied to the output from the previous time step. The internal state parameter is employed to compute the output. Brownlee (2017)The primary components of the LSTM memory cell are the gateways. These are weighted mechanisms that regulate the movement of information within the cell. Three gateways are present. The neglect gateway determines the information to be discarded from the cell. The intake gateway defines which input values update the memory state. The egress gateway dictates the output based on the input and the cell's memory.

Zhou and Zhu (2014)The LM-BP neural network is demonstrated to achieve a faster convergence than the traditional BP algorithm, offering good predictive ability and great accuracy. Wang, Zentner, Pedroni and Zio (2017)The calculation of fragility curves is a very important element in risk analysis and by developing an artificial neural network-based statistical metamodel, an approach that can perform seismic fragility analysis in a shorter time is developed.

Balkasoglu and Yildirim (2018)In the study carried out utilizing with LSTM approach within the framework of recurrent artificial neural networks, the incorporation of time interval modeling and data focused on fault lines could potentially yield more favorable outcomes.Yildirim (2018)It is noted that the suggested wavelet-oriented stratum within the deep bidirectional LSTM network framework, referred to as DBLSTM-WS, for the categorization of electrocardiogram (ECG) signals, substantially enhances the recognition efficacy compared to conventional networks.

Sivaiah Bellamkonda et al. (2021)It is seen that the results obtained in terms of LSTM are superior compared to FFNN results. They believe that if the current method is combined with deep learning methods, the results can be obtained precisely.

Wang, Guo, Yu and Li (2017)Indicates that the LSTM model with two-dimensional input established in the study can identify and utilize spatio-temporal connections among earthquakes, resulting in improved predictions compared to previous methods. Al Banna et al. (2021)In a temporal sequence analysis conducted on earthquakes in Bangladesh, when scientists reviewed prior research, they indicated that among the available deep learning algorithms, LSTM yielded the most promising outcomes.

Berhich et al. (2021)Optimal hyperparameter values of the models are obtained by the grid search technique. The application of the LSTM model to the same data set was compared with the multilayer neural network model. Looking at the final results, the improved LSTM model provides effective prediction and achieves better performance compared to others. This improved earthquake prediction model was developed by Berhich et al. (2022) and it can be said that it gives better performance and better results.

Nicolis et al. (2021)There is a notion that the deep learning model suggested in the research could be advanced through the incorporation of additional external factors like seismic depth, crustal movement, proximity to major faults, and other geological parameters. This enhancement is particularly relevant for predicting main earthquakes, where ETAS models typically encounter limitations.

Karci and Sahin (2022)The relationship between earthquake data and earthquake time was examined with a neural network approach such as Long-Short-Term Memory (LSTM). In the forecasts for the 15-day and 30-day periods, the LSTM model best predicted the highest and lowest earthquake magnitudes. Wang et al. (2023)The Random Forest (RF) method is the best way for categorising significant earthquake occurrences, and the LSTM method offers a reliable estimate of earthquake magnitude.Berhich et al. (2023)In the study using earthquake datasets, which is a seismically active region, using Long-Short Term Memory (LSTM), it is stated that their proposed model performs significantly better than other methods.

The research covering the most advanced technologies and methods for earthquake prediction are examined. When the studies in Table1 are analysed, it is seen that although the studies do not provide definite results, they constitute an important basis for making more precise predictions in the future in regions under earthquake risk. It can say that artificial intelligence-based methods, especially deep learning models such as LSTM and Random Forest, are very effective in estimating earthquake magnitude and time. The development and use of such technologies may contribute to more accurate and reliable earthquake predictions in the future. Due to the complexity and uncertainty of studies on earthquake prediction, it is, unfortunately, insufficient to obtain definitive results with today's technology. Most research still has limitations in estimating the exact time and location of earthquakes. Advanced neural network models and deep learning techniques may require large amounts of data, and obtaining such data can be difficult and costly. In addition, the learning processes of some models can be quite time-consuming and require complex calculations. Different methods and models achieve various successes in the field of earthquake prediction. Artificial neural networks, deep learning, and statistical methods play an important role in earthquake prediction, and good results can be obtained in earthquake prediction in the future with the development of technology, large databases, and seismic technologies, as well as multi-disciplinary studies in earthquake prediction.

Paper	Dataset	Target	Method	Predicted variables	Range of magnitude	Type of prediction
Balkasoglu	Turkiye	Earthquake prediction	LSTM	Time,Magnitude,Location,	Magnitude $\bar{>}$2.0	Regression
Karcı	Turkiye	Earthquake prediction	LSTM,DT,RF	Time,Magnitude,Location, Depth	Magnitude $\bar{>}$3.5	Regression
Al Banna	Bangladesh	Earthquake Prediction	Attention and Bi-LSTM	Earthquake or non-earthquake occurrence	< 7 Magnitude < 7.5	Calssification
Berhich	Japan, Turkey, and Morocco	Earthquake Prediction	LSTM - GRU	Magnitude occurrence, location cluster,time	Magnitude< 5.0	Clustering and Regression
Sivaiah	None	Earthquake Prediction	LSTM	Time, Magnitude, Location, Depth, Magnitude	Magnitude< 8.0	Calssification
Wang	China	Earthquake Prediction	Bi-LSTM	spatio-temporal correlations among earthquakes	Magnitude< 4.5	Classification
Mallouhy	Nothern California	Earthquake Prediction	SMV, RF, AdBoost,CART	Major earthquake prediction, short term, negative and positive earthquake	Magnitude> 5.0	classification
Yaghmae	Chi-Chi, Taiwan	Earthquake Prediction	ANN, PNN, GRNN	site classification	Magnitude$\bar{>}7,6$	Classificationn
Khawaja	Hindukush	Earthquake prediction	SVR and HNN	Earthquake or Non-Earthquake Occurrence	Magnitude> 5	Classification
Asim	Cyprus	Seismic Activity Rredictions	RF, ANN, and SVM	Earthquake or Non-Earthquake Occurrence	Magnitude $3.0, 3.5, 4.0, 4.5$	Classification
Berhich	Japan, Red Sea.	Earthquake Prediction	LSTM	Time, Magnitude,Location	Magnitude > 6	Regression
Nicolis	Chile	Location of Seismic Prediction	LSTM, CNN	Spatio-Temporal Estimation module	Magnitude > 4	Classification
Reyes	Chile	predict earthquakes	ANN, KNN, SVM, K-means	Seismic activities	Undefined	Time series
Xi	China	Earthquake Predict	RF,LSTM	Predict Large Earthquakes	Magnitude > 6	Clustering and Regression
Vardaan	Nepal,India.	Earthquake Prediction	RNN,LSTM, FFNN	Time,Magnitude,Location, depth	Magnitude> 3	Regression
Gul	Turkey 1975-2016	Damage assessment	LM- ANN Time		Magnitude> 5	Time series
Agha	Bam in 2003	Damage assessment	BPNN	Human Loss Estimation	Undefined	Time series
Malkocoğlu	NOAA 2021	Damage Estimation	ANN, PSO	natural phenomena	Magnitude $\bar{>}$3	Classification
Xing	China in 1970-2015	Damage assessment	RW SVM, BPNN	Casualty Fundamental	Undefined	Classification
Yuan	U.S Geological Survey(1900-2019	Earthquake Magnitude Prediction	K-means	A Seismic Prediction	Undefined	Clustering
Natarajan	Iran 2017	Damage assessment	RF, DT	Prediction Damage	Undefined	Regression

Table 1: Comparative analysis of earthquake prediction research in existing literature

3 Methodology

Deep learning architecture was used for an analytical research to create an earthquake prediction model with real earthquake data. Big data infrastructures are needed for research studies on earthquake prediction and similar topics. Big data analysis processes, which include information technology processes, must be carried out by going through various stages starting from pre-analysis to the last step in order to provide good results in the data analysis process. In these processes, after it is accepted that each previous step meets the necessary criteria, the next step is taken. In this study, the research steps are as follows: data collection, data preprocessing, data transformation, data selection, data modeling, evaluation, predictions, and results. Within the scope of this study, the analysis of the deep learning processes of earthquake prediction models was carried out within the framework of six stages.

3.1 Methodology Stages

3.1.1 Data Collection

In analysis studies, robust data collection is vital for reliable results. The Earthquake Query System from Kandilli Observatory and Earthquake Research Institute provides crucial earthquake prediction model components, including time, location, depth, and magnitude. The National Earthquake Monitoring Center manages data operations, converting, evaluating, archiving, and distributing records. Collected earthquake data from Turkey (1900-2018) are stored in ".txt" files, creating a publicly accessible resource for analysis.

3.1.2 Data Pre-Processing

Data preprocessing improves data quality for analysis or modeling, enhancing accuracy. It involves cleaning, correcting missing info, and proper formatting. In earthquake research, data from the National Earthquake Monitoring Center is used, cleaned, and filtered based on specific criteria. This ensures no missing, duplicate, or biased values. The process clarifies earthquake details like time, location, depth, and magnitude, excluding other variables.

3.1.3 Data Transformation

Data transformation, a part of preprocessing, readies data for analysis or modeling by enhancing understandability and analyzability. Date format for earthquake occurrence time is converted. Earthquake data is stored in (.xls) files with related information.

Taking logarithms of features like latitude, longitude, depth, and magnitude normalizes their ranges and balances value differences across scales. This aids machine learning model training by mitigating dimensional disparities. Logarithms compress values, reducing the impact of large ones for balanced training. It also captures non-linear relationships, improving modeling of such relationships.

3.1.4 Data Selection

Data selection is the process of identifying the data set to be used for an analysis or modeling study and removing redundant or less effective data. This is an important step to increase the efficiency of the analysis process and improve the performance of the model. Data The analysis process was completed with these criteria under five headings: Turkey earthquake date, 1900 - 2018, latitude; 34.00 - 43.00, longitude; 25.00 - 46.00, Depth; 0-500 km, Magnitude "4.0 - 9.0 Mw", Shake type. The date and time, latitude, longitude, depth and magnitude (xM) of the earthquake, as well as other information to be used in the earthquake prediction model, were removed from the data. The date format of the earthquake date data was changed to year, month, day, hour, minute and second.

Time Stamp: The moment the event happened. The measurement of time is in milliseconds. It denotes the "origin" time or the moment the earthquake starts to rupture.

Latitude: Expressed in degrees. An earthquake initiates rupture at a hypocenter, defined by a location on the earth's surface (epicenter) and a depth beneath this point (focal depth). Latitude represents the number of degrees north (N) or south (S) of the equator and ranges from 0 at the equator to 90 at the poles. Longitude ranges from

0 at the Prime Meridian to 180, with E or W indicating the direction from the Prime Meridian.

Longitude: Expressed in degrees of longitude. The longitude represents the count of degrees east (E) or west (W) of the prime meridian, which passes through Greenwich, England. Longitude ranges from 0 at Greenwich to 180, with the E or W indicating the direction from Greenwich.

Magnitude: Earthquake magnitude is an assessment of the scale of an earthquake at its origin. It follows a logarithmic scale.

Depth: Event depth in kilometers, where earthquake rupture starts. This depth is referenced to mean sea-level or seismic station elevation, determined by earthquake location method and seismic network.

3.1.5 Data Modeling

Data modeling establishes the groundwork for analytics or AI applications, crafting mathematical and statistical models for data comprehension. It uncovers patterns, relationships, and data behavior. This study amalgamates feature selection and time series forecasting to optimize system resource workload prediction. Data mining algorithms learn patterns from historical data to predict outcomes of unknown instances. To test and compare, the chosen model is Long-Short-Term Memory (LSTM).

LSTM excels in analyzing time series data and detecting intricate patterns through its memory cell and gates. It finds application in various domains like natural language processing and voice recognition. Its proficiency in grasping long-term dependencies makes it crucial in deep learning. Post LSTM estimation, outcomes are compared with GRU, Bidirectional LSTM, and RNN to determine the best-performing model.

3.1.6 Data Evaluation

Model evaluation uses performance metrics to assess machine learning model strengths/weaknesses. Scores come from test examples and predictions, aiding feature selection during training. Shcherbakov et al. (2013) emphasize diverse metrics for time series forecast error quantification. While no predefined criteria exist, RMSE, MAE, and MAPE are recommended for evaluation.

RMSE is a metric that measures the spread of forecast errors and indicates how accurate the model's predictions are. A low RMSE value means good fit, but outliers can affect it. MAE refers to the sum of the absolute differences between actual and predicted values and is robust to outliers. A low MAE value indicates good model performance. MAPE is the average absolute value of the prediction errors as a percentage and a low MAPE value indicates a good fit model.

3.1.7 Predictions

The Predictions section includes the evaluation of selected models with invisible data. It examines which model gives better results over MSE, RMSE and MAE results of LSTM, RNN, GRU, BiLSTM models.

4 Design Specification

This project follows a three-layer architecture: Data Persistent Tier, Business Logic Tier, and Client Tier. These layers collaborate to execute the analysis steps. The first layer handles persistent data, the second implements business logic, and the third serves as the client interface. The project utilizes popular Python libraries such as Numpy, Pandas, Scikit-learn, and TensorFlow for AI analysis.

Data Persistent Tier: This layer is where data is stored and processed. Using persistent storage solutions such as databases or file systems, data is stored there and used by the business logic layer as needed.

Business Logic Tier: In this layer, business logic and data modeling steps are performed. The data is processed and associated with the specified models. The LSTM (Long Short Term Memory) layer, used especially for prediction, is a combination of bidirectional and dense layers. LSTM is a common deep-learning structure for analyzing time series and sequential data. LSTM was preferred for studies such as earthquake prediction involving a long time period. The results obtained are evaluated by comparing them with the outputs of models such as RNN, GRU, and BiLSTM. Performance analysis is done based on MSE, RMSE, and MAE results.

Mean Squared Error quantifies the average of the squared disparities between the actual and forecasted values within the dataset. It gauges the dispersion of the residuals, reflecting the extent of variability present in the predictions.

Root Mean Square Error, is a metric that gauges the accuracy of predictions by calculating the square root of the average of squared differences between predicted and actual values within a dataset. A lower RMSE indicates a better fit of the model to the data and allows for easy comparison with other models. However, it can be sensitive to outliers or large values.

Mean Absolute Error, quantifies prediction accuracy by summing up the absolute disparities between predicted and actual values. It treats deviations without considering their direction, making it resistant to the impact of outliers. A smaller MAE signifies effective model performance.

Client Tier: This layer is for the end user. Users access the system through this layer and make the necessary requests to generate forecasts. In this layer, the process of presenting the estimation results in a more understandable, visual and similar way for users is carried out.

4.1 Recurrent Neural Network

In conventional neural networks, input and output are disjoint, whereas in RNN, the output relies on prior elements in the sequence. Furthermore, recurrent networks reuse parameters across all layers. Unlike feedforward networks where each node has distinct weights, RNN employs uniform weights within each layer. During gradient descent, RNN adjusts weights and biases autonomously to minimize loss.

A recurrent neural network (RNN) is one of two broad types of artificial neural networks characterized by the direction of information flow between its layers. Unlike a one-way feedforward neural network, it is a bi-directional neural network, that is, it allows the output from some nodes to affect subsequent inputs to the same nodes.

4.2 Long-Short Term Memory

LSTM (Long Short Term Memory) is a special type of RNN developed by Hochreiter and Schmidhuber (1997) to address addiction problems. All recurrent neural networks consist of recurrent modules that are arranged in a chain. While in standard RNNs this module has a simple "tanh" activation function, LSTMs have a more complex structure. In LSTMs, the recurrent module consists of four specialized layers and realizes the interaction more efficiently. Unlike standard RNNs, the LSTM module is designed to learn and process long-term dependencies more effectively. The Figure 1 shows a repeating LSTM module structure with four layers.

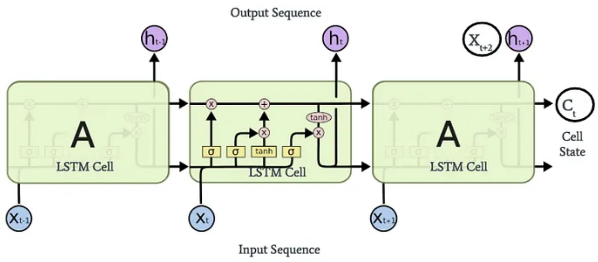

Figure 1: A Long-Short-Term Memory ModuleJanardhanan and Barrett (2017)

4.3 Gated Recurrent Unit

Cho et al. (2014)It is an enhanced version of the GRU standard recurrent neural network developed by GRUs were developed to solve the vanishing gradient problem. GRU can also be considered as a variation of LSTM because both are similarly designed network structures and in some cases can yield equally successful results.Kostadinov (2017) The main purpose of both is the same. LSTMs and GRUs were devised to overcome the challenges posed by limited short-term memory. These models incorporate internal "gate" mechanisms that have the ability to manage and control the information flow within the network..Phi (2018) However, GRUs have fewer parameters than LSTM because they do not have an exit port. Figure 2 Gates of GRU and LSTM are shown.

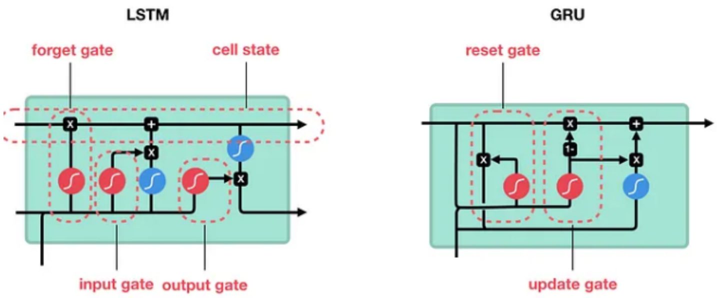

Figure 2: Gates of GRU and LSTMPhi (2018)

Demirci and Karatli (2023) These gates learn to prioritize essential training data and can sift through extensive sequences to make accurate predictions. Most advanced results in RNNs are achieved through these networks. Phi (2018)LSTMs and GRUs are applicable in speech recognition, speech synthesis, and text generation, and they can also generate subtitles for videos. Weiss et al. (2018) showed that LSTM gave better results than GRU in their study. The reason for this is that while LSTMs can easily do unlimited counting, the GRU cannot. Because of this, the GRU fails to learn the simple languages that can be learned by LSTM.Gruber and Jockisch (2020) study they showed that GRUs performed better on smaller and less frequent datasets.

4.4 Bidirectional Long Short-Term Memory (BLSTM)

LSTMs, a novel form of RNN, emerged to address the challenge of Vanishing-Gradient and manage extended connections stemming from limited-term memory. In contrast to RNNs, LSTMs incorporate memory cells, along with interlinked gates designed in a specific manner. Donges (2018) These gates are referred to as Input Gate, Reset Gate, and Output Gate, yielding a value within the range of 0 to 1. This outcome determines what information to retain, discard, select, or gather from the data. In a time-dependent context, incoming data sequences merge with memory containing past outputs, generating fresh outputs that are stored within memory.

Sahingoz et al. (2020) Bidirectional LSTM processes input data in two directions: one from past to future and the other from future to past. This bidirectional approach allows the LSTM to capture information from both directions, preserving insights from both past and future contexts. Moreover, Bidirectional LSTM effectively integrates the two hidden states, resulting in combined states that converge harmoniously.

5 Implementation

This project was conducted to develop an earthquake prediction model. In this application project, an earthquake prediction model was created using seismic data of the regions and deep learning architectures. Various libraries of Python programming language were used for the deep learning process.

This research aims to obtain the best result on time series in earthquake prediction. In this process, Long-Short Term Memory (LSTM), one of the deep learning architectures, will be used. The analysis process is completed using various libraries of the Python programming language. In the implementation process, libraries such as pandas, numpy, seaborn, matplotlip data visualization, TensorFlow, Keras, scikitlearn and ydata_profiling were used. The results of the analysis will be compared with the results of other architectures such as RNN, Bidirectional LSTM, Gated Recurrent Unit and Long-Short Term Memory (LSTM), one of the deep learning architectures. The implementation process consists of selecting the tracking data, loading it into the system, cleaning, transforming, modeling, and model results.

Earthquakes listed according to the specified criteria were saved in files with ".txt" extension. Earthquake catalog data were converted to ".xls" format. The libraries to be used in the analysis process were imported into Python. TensorFlow was used in the application project, but building deep learning models can be difficult. Therefore, Keras was preferred. Keras is a Python library that makes it easy to create deep learning models.

Configuration files were included on my system to load configuration information for the application project. Within the project, configuration information is typically used to contain the variables, settings and configurations required for the project to run. Two different directories called "models_path" and "reports_path" were created to store and access the relevant directory paths for use in different parts of the project. These directories make the file management and storage process more convenient.

In the second step of the implementation project, the file containing the earthquake data was uploaded to our system. Here, the original file name was merged into the raw data directory path in a fixed variable. The data of the original file consists of 6574 rows and 10 columns. To avoid corruption of the original data, a copy of the original data was created. This preserved the original data while working on the data frames.

Next, the file is filtered to hold only data relevant to the analysis. This involves creating a new data frame with columns for 'time', 'Latitude', 'Longitude', 'Depth' and 'Magnitude'. This simplifies the data frame for focused analysis. Next, the 'Time' column is formatted as "'%Y-%m-%d %H:%M:%S'", making it suitable for time series or specific timestamp analysis. This column acts as an index facilitating quick access to the data corresponding to a specific timestamp.

Subsequently, data is sorted based on the "index," ensuring an orderly time series with ascending timestamps. This sorting aids in comprehending trends and changes over time. The index is converted to a DateTime object, enhancing convenience and accuracy for time series analysis. DateTime objects are a widely-used data type in Python for handling date and time values, streamlining temporal analyses.

ProfileReport is used to analyze datasets, providing statistics and insights for quick comprehension and valuable insights. This tool automatically generates analyses, including graphical representations, for "Latitude, Longitude, Depth, Magnitude" values.

Logarithmic data transformation involves taking the natural logarithm of data, particularly beneficial for managing large value ranges. This enhances distribution, mitigates variability, and improves machine learning model performance. Logarithmic transformation was applied to earthquake data. The dataset is split into train and test files, with train_data comprising about 90% of the dataset and test_data encompassing approximately 10%.

A TimeSeriesGenerator is used to organize training data into time steps for machine learning. It groups time series data and feeds it to the model during training. Loops in TimeSeriesGenerator generate training and test data, smoothing input and output data within specified periods. This data is inserted into arrays like train_x, train_y, test_x, and test_y. The input sequences are reshaped into a three-dimensional structure, as expected by TimeSeriesGenerator, with data samples treated as matrices. This structure organizes the number of matrices, features, and input length according to defined values.

Data reconstruction ensures proper dimensions for training and test data, enabling successful model training and evaluation. The input array sizes (5988, 4, 12) for training and (572, 4, 12) for test indicate suitable data shaping for the model. The chosen deep learning algorithm is LSTM, compared with RNN, BiLSTM, and GRU results.

The LSTM model architecture includes sequential building using the "Sequential" function. It comprises two LSTM layers: the first with 64 cells using tanh activation, and the second with 32 cells also using tanh activation. Additionally, two dense layers are included, with 32 neurons using ELU activation in the first layer, and 16 neurons using GELU activation in the second layer. The output layer matches the columns in the training data's output array.The "loss" and "val_loss" values from training are vital metrics for assessing model performance. They measure the model's fit to training data and its generalization to validation data.

MSE, RMSE, and MAE will assess model performance. Lower values imply better predictions closer to actual data. Smaller metrics suggest more successful predictions, but their acceptability depends on project needs and dataset characteristics.

6 Evaluation

In this research project, an evaluation on an earthquake prediction model was made with real earthquake data and algorithms such as Long-Short Term Memory, Recurrent Neural Network, Gated Recurrent Units, Bidirectional Long-Short Term Memory. In order to compare the model results applied for each algorithm, all results were compared with the widely used Root Mean Square Error. At this stage, all models are compared in order and it is stated which model provides the best result. In the table in the last part of this title, the results of the research conducted on this subject and the results of the application research are compared.

6.1 Long-Short Term Memory Result

This model includes Long-Short Term Memory layers. It is structured as a ”sequential” model consisting of five layers. There are a total of 33,780 parameters in the model. These parameters represent the learnable weights of the model.

The loss value obtained in training the model is 21.74 and the validation loss (val_loss) is 182.13. These results show the performance of the model on the training data. The fact that the validation loss is greater than the training loss indicates the risk of overfitting the model.

Considering the calculated performance values, the mean squared error (MSE), which is used to measure how far the true values are, is calculated as 195.54. Root mean square error (RMSE), which is used to calculate the average magnitude of prediction errors, is calculated as 13.98. The mean absolute error (MAE) used to calculate the mean absolute magnitude of the prediction errors was calculated as 5.71.

In general, based on the training results of the model, it is seen that the predictions contain a certain margin of error compared to the actual values. However, the high validation loss of the model and the significant MAE value may suggest that the model can be improved to achieve better performance. These improvements can be achieved by changing the model structure, using more data or hyperparameter adjustments. Figure 3 is analyzed, and it is observed that among the variables in the model, the predictions that best explain the actual values are concentrated on the ”depth” variable.

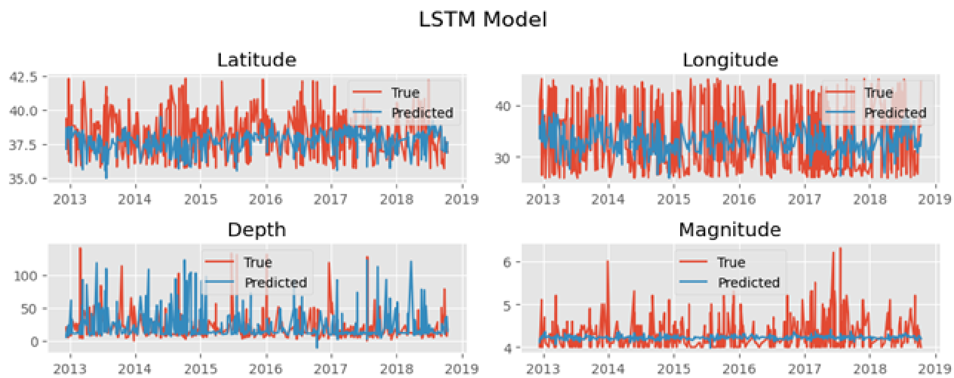

Figure 3: Long-Short Term Memory Result

6.2 Recurrent Neural Network Result

This model contains layers of Recurrent neural network. This model is also structured as a "sequential" model. There are a total of 9684 parameters in the model and these parameters represent the learning ability of the model

The loss value obtained in training the model is 152.00 and the validation loss (val_loss) is 138.57. These results show the performance of the model on the training data. However, the fact that the training loss is high and the validation loss is similarly high may suggest that the model may be a poor fit to the data.

Looking at the calculated performance values, the mean squared error (MSE), which is used to measure how far away the true values are, was calculated as 95.13. Root mean square error (RMSE), which is used to calculate the average magnitude of prediction errors, was calculated as 9.75. The mean absolute error (MAE) used to calculate the mean absolute magnitude of the prediction errors was calculated as 4.70. According to these results, it is seen that the predictions of the model are still high compared to the actual values. When the performance metrics are analyzed, it is observed that the deviation and errors in the predictions of the model are evaluated.

Consequently, different approaches and methods can be tried to further improve and optimize the performance of the model. Using more data, hyperparameter adjustments or examining different layer structures can improve the overall performance of the model."Figure 4 is analyzed, it is seen that the predictions that best explain the actual values of the variables in the model are concentrated in the "depth" variable.

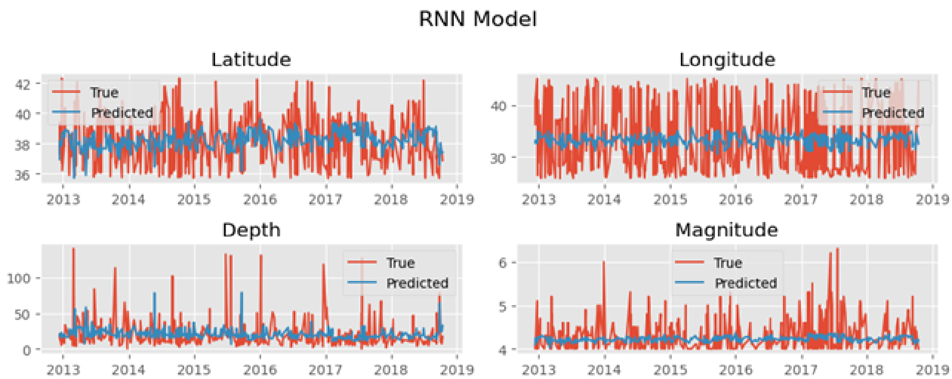

Figure 4: Recurrent Neural Network Result

6.3 Gated Recurrent Units Result

This model includes Gated Recurrent Units (GRU) layers. This model is also structured as a "sequential" model. There are a total of 26,036 parameters in the model and these parameters represent the learning capability of the model

The loss value obtained in training the model is 43.73 and the validation loss (val_loss) is 178.73. The fact that the model training loss and validation loss are also high may suggest that the model may be a poor fit for the data.

Considering the calculated performance values, the mean squared error (MSE), which is used to measure how far away the true values are, is calculated as 129.07. Root mean square error (RMSE), which is used to calculate the average magnitude of prediction errors, was calculated as 11.36. The mean absolute error (MAE) used to calculate the mean absolute magnitude of the prediction errors was calculated as 4.99.

As a result, the use of "GRU" layers in this model may help the predictions to give better results. However, the high loss values and MAE values may suggest that the model needs to be improved to achieve better performance. These improvements can be achieved by changing the model structure, using more data or hyperparameter adjustments."Figure 5 is analyzed, it is seen that the predictions that best explain the actual values of the variables in the model are concentrated in the "depth" variable.

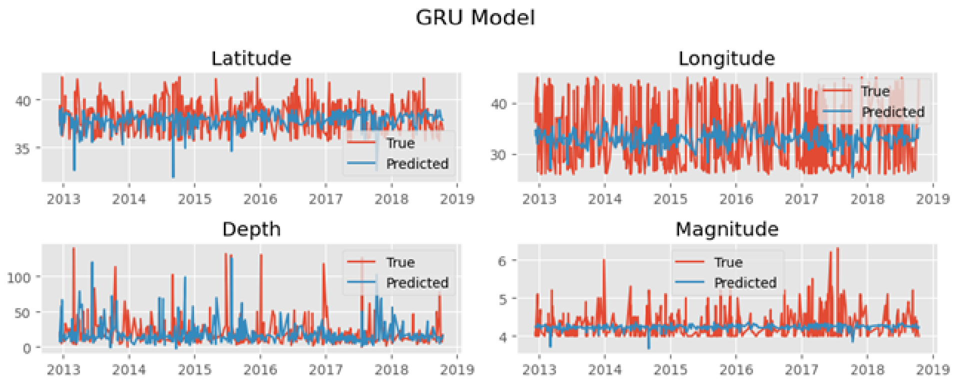

Figure 5: Gated Recurrent Units Result

6.4 Bidirectional LSTM (Bilstm) Result

This model includes the Bidirectional LSTM Model (GRU) layers. This model is also structured as a "sequential" model. The model has a total of 83,316 learnable parameters.

The loss value obtained in training the model is 5.43 and the validation loss (val_loss) is 165.99. The fact that the training loss is low but the validation loss is high may indicate that the model has the risk of overfitting.

When we look at the values seen in the last epoch of the training process, it is seen that the training loss is 5.43 and the validation loss is 165.99. MSE value is 156.49, RMSE value is 12.51 and MAE value is 5.25. When we look at the performance metrics, we see that the mean square error (MSE) value, which measures how far the predictions are from the true values, is high. Similarly, the root mean square error (RMSE) and the mean absolute error (MAE) values are not low. This may suggest that the model's predictions need further improvement in terms of accuracy.

In conclusion, the training results and performance metrics of the model show that there is room for further tuning and improvement. Methods such as data manipulation, hyperparameter adjustment or different model structures can be considered for the model to gain a better generalization capability. "Figure 6 is analyzed, it is seen that the predictions that best explain the actual values of the variables in the model are concentrated in the "depth" variable.

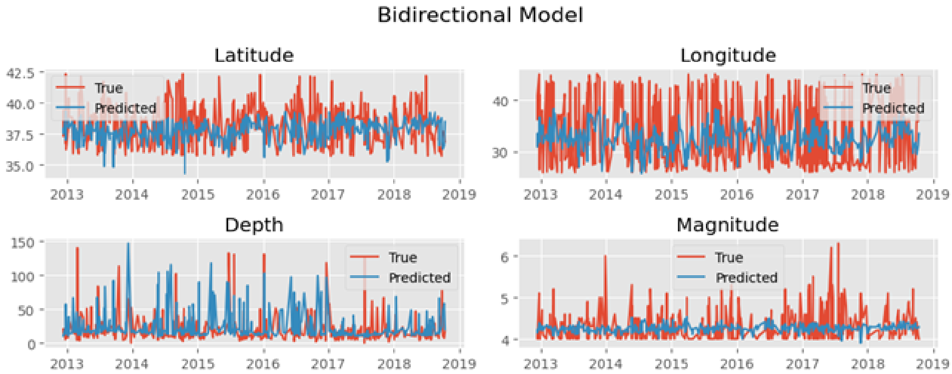

Figure 6: Bidirectional Long-Short Term Memory Result

In cases where all of the study project's findings are looked at general; In the results of the LSTM model, which is the main model chosen for our research, low training loss and high validation loss may indicate overfitting, and high MSE, RMSE and MAE values may indicate that the predictions are generally far from the true values. In the RNN model results, higher training and validation loss may increase the likelihood of overfitting, and lower MSE, RMSE, and MAE values may indicate better predictions than other models. The GRU model, on the other hand, may be more likely to generalize with moderate loss of training and validation, and lower MSE, RMSE, and MAE values compared to other models indicate generally better predictions. Considering the BiLSTM model, low training and validation loss can be a good generalization indicator and since MSE, RMSE and MAE values are lower than other models, it can be said that the estimations are more sensitive and close to the true values.

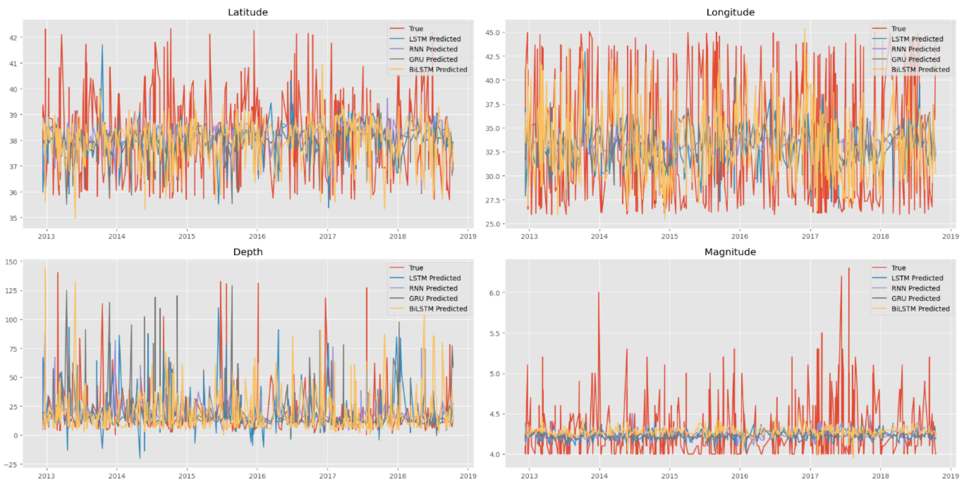

Figure 7: All Models Result

In general, the model with the best results among the examined models appears to be "RNN". Lower training and validation loss, lower error metrics, and better predictions show that this model better fits the data. On the other hand, the "BiLSTM" model also performs quite well and the predictions are generally more accurate than other models. In the LSTM model, on the other hand, it is understood that there is a high risk of overfitting and the predictions are often far from the true values. The GRU model, on the other hand, performs more balanced than other models, but it can still reduce prediction errors. "Figure 7 it is seen that the "depth" variable provides the best results in the process of best "estimating" the actual data in the research project.

The results obtained in similar studies are given in Table 2. The models used by the researchers, the regions covered by the data sets and some metric results are also included in this table. Considering the results of the researchers in the researches in Table 2,Wang et al. (2023) LSTM model applied in the Sichuan-Yunnan region of China. The MSE and RMSE values are high, indicating that the predictions deviate more from the actual values.Berhich et al. (2023)LSTM model applied to data from Turkey. A low MSE value may indicate that the predictions are close to the actual values. RNN model used for the same data set. The MSE value is low, but the MAE value is higher, indicating that the predictions may deviate more on average. Cetina (2014) research in Turkey, high R2 value shows that the model explains the data well, while low MSE and RMSE values indicate that the estimates are close to the real values. These results demonstrate the strong performance of the model.

Paper	Dataset	Model	R2	MSE	RMSE	MAE
Omer	Turkey	LSTM		128.69	11.344	5.042
Omer	Turkey	RNN		103.29	10.16	4.88
Omer	Turkey	GRU		142.51	11.93	5.12
Omer	Turkey	BiLSTM		100.2	10.01	4.51
Karcı and Sahin	Turkey	LSTM	0.506	0.151	0.367	0.283
Wang et al.	China	LSTM		0.2289	0.4784	0.4193
Berhich et al.(2023)	Turkey	LSTM		0.0151		0.0941
Berhich et al.(2023)	Turkey	RNN		0.0157		0.11
Berhich et al.(2023)	Japan	LSTM		0.0103	0.1014	0.0754
Berhich et al.(2023)	Northern Red Sea	LSTM		0.0139	0.11	0.0886
Cetina	Turkey	ANN	98.5		0.13	0.11
Malkoçoğlu et al.	Turkey	ANN		0.95	0.96	
Cui et al.	China	XGBoost	0.5091	0.3469		0.4454
Aghamohammadi et al.	Western Iran	ANN	0.9733		0.111	
Wen et al.	Japan	Deep Neural Network	0.876	0.106		0.251
Natarajan et al.	Iran	RandomForest	0.04		2.92	0.11
Al Banna et al	Bangladesh	LSTM		1.5579	1.24818	

Table 2: Comparative analysis of earthquake prediction research in existing literature

Karci and Sahin (2022)LSTM model applied on Türkiye data by A high R2 value indicates that the model explains the data well. Malkocoglu et al. (2022)ANN model applied to data from Turkey. High R2 and MSE values may indicate that the model explains the data well, but a high RMSE value may indicate that the predictions show more deviations compared to the actual values.

Cui et al. (2021), XGBoost model applied to Chinese data. A moderate R2 value and MSE value indicate that the predictions are generally good, but there are still deviations with RMSE and MAE values.Berhich et al. (2023)LSTM model applied to data from Japan. A low R2 value may indicate that the model does not explain the data well. MSE and RMSE values indicate that the predictions deviate from the actual values.Wen et al. (n.d.)Deep Neural Network model applied to data from Japan. A high R2 value indicates that the model explains the data well.

Aghamohammadi et al. (2013)ANN model applied in the Laleh Valley of Iran. The high R2 value indicates that the model explains the data very well.Natarajan et al. (2023) Random Forest model applied to Iranian data. A low R2 value and a high MSE value indicate that the model has difficulty explaining the data and there are large deviations in the predictions.

Al Banna et al. (2021) LSTM model applied on Bangladesh data. High MSE and RMSE values indicate that the predictions contain large deviations compared to the actual values.Berhich et al. (2023)LSTM model applied to Northern Red Sea data. A low R2 value may indicate that the model does not explain the data well, and MSE and RMSE values may indicate that the predictions deviate from the actual values. Artificial Neural Network (ANN) model applied on Turkey data. A high R2 value indicates that the model explains the data very well.

Here, the performance of different models on different datasets is analysed. Various deep learning models such as LSTM, RNN, GRU and BiLSTM are used and evaluated with metrics such as R2 score, MSE, RMSE and MAE. The results show that there are significant variations between different data sets and models. Model performance varies depending on the characteristics of the data set, the characteristics of the model and the choice of metrics used. High R2 values and low MSE, RMSE and MAE values indicate better forecasting performance, while low R2 values and high error metrics indicate that the model can be improved. By highlighting the impact of different model and data combinations on forecasting, this analysis can provide guidance for future similar studies and forecasting projects.

7 Conclusion and Future Work

This project aimed to develop an earthquake prediction model using earthquake data from Turkey. The objective was to forecast future earthquakes by leveraging a dataset containing critical information such as time, latitude, longitude, depth, and magnitude of past earthquakes. This approach, rooted in big data analysis, has been widely explored in earthquake prediction research. The selection of the Long-Short Term Memory (LSTM) deep learning technique for the model was motivated by its capacity to uncover temporal patterns and relationships.

Examining the outcomes of the chosen model, a notable trend emerged where the training loss was low, while the validation loss was relatively high. This discrepancy suggests that the model excelled at fitting the training data but struggled to generalize to new data. To rectify this, strategies like augmenting the dataset, refining model hyperparameters, and optimizing the model's architecture can be employed.

To enhance the efficacy of earthquake predictions in the future, a primary focus should be placed on amassing comprehensive and dependable data. Enlarging and updating the databases can enhance the model's training capabilities. Additionally, integrating diverse deep learning and conventional statistical methods could heighten prediction accuracy. Exploring the incorporation of earthquake precursors, geographical factors, and varied parameters into the model may also contribute to improved forecasts. In the realm of earthquake prediction studies, collaborating across scientific disciplines and designing innovative model infrastructures holds the promise of yielding superior outcomes.

Acknowledgement

I would like to thank my research supervisor Mr. Abdul Shahid for his support during my coursework.

References

Aghamohammadi, H., Mesgari, M., Mansourian, A. and Molaei, D. (2013). Seismic human loss estimation for an earthquake disaster using neural network, *International Journal of Environmental Science and Technology* **10**: 931–939.

Al Banna, M. H., Ghosh, T., Al Nahian, M. J., Taher, K. A., Kaiser, M. S., Mahmud, M., Hossain, M. S. and Andersson, K. (2021). Attention-based bi-directional long-short term memory network for earthquake prediction, *IEEE Access* **9**: 56589–56603.

Alarifi, A. S., Alarifi, N. S. and Al-Humidan, S. (2012). Earthquakes magnitude prediction using artificial neural network in northern red sea area, *Journal of King Saud University-Science* **24**(4): 301–313.

Altun, F. (2018). Afetlerin ekonomik ve sosyal etkileri: Turkiye ornegi uzerinden bir degerlendirme, *Sosyal Calisma Dergisi* **2**(1): 1–15.

Asencio-Cortes, G., Martinez-Alvarez, F., Morales-Esteban, A., Reyes, J. and Troncoso, A. (2015). Improving earthquake prediction with principal component analysis: Application to chile, *Hybrid Artificial Intelligent Systems: 10th International Conference, HAIS 2015, Bilbao, Spain, June 22-24, 2015, Proceedings 10*, Springer, pp. 393–404.

Asim, K. M., Moustafa, S. S., Niaz, I. A., Elawadi, E. A., Iqbal, T. and Martinez-Alvarez, F. (2020). Seismicity analysis and machine learning models for short-term low magnitude seismic activity predictions in cyprus, *Soil Dynamics and Earthquake Engineering* **130**: 105932.

Azam, F., Sharif, M., Yasmin, M. and Mohsin, S. (2014). Artificial intelligence based techniques for earthquake prediction: A review, *Sci Int* **26**(4): 1495–1502.

Balkasoglu, S. and Yildirim, O. (2018). Regional using a deep ukvh network model earthquake estimation, *2018 Innovations in Intelligent Systems and Applications Conference (ASYU)*, IEEE, pp. 1–5.

Berhich, A., Belouadha, F.-Z. and Kabbaj, M. I. (2021). Lstm-based earthquake prediction: Enhanced time feature and data representation, *International Journal of High-Performance Systems Architecture* **10**(1): 1–11.

Berhich, A., Belouadha, F.-Z. and Kabbaj, M. I. (2022). A location-dependent earthquake prediction using recurrent neural network algorithms, *Soil Dynamics and Earthquake Engineering* **161**: 107389.

Berhich, A., Belouadha, F.-Z. and Kabbaj, M. I. (2023). An attention-based lstm network for large earthquake prediction, *Soil Dynamics and Earthquake Engineering* **165**: 107663.

Bikce, M. (2017). Taken and need to take precautions in turkey's earthquakes, *International Journal of Engineering Research and Development* **9**(2): 24–31.

Bilen, M., Isık, A. H. and Yiğit, T. (2015). A hybrid artificial neural network-genetic algorithm approach for classification of microarray data, *2015 23nd Signal Processing and Communications Applications Conference (SIU)*, IEEE, pp. 339–342.

Bodri, B. (2001). A neural-network model for earthquake occurrence, *Journal of Geodynamics* **32**(3): 289–310.

Brownlee, J. (2017). *Long short-term memory networks with python: develop sequence prediction models with deep learning*, Machine Learning Mastery.

Buyukkaracigan, N. (2016). Turkiye'de yerel y onetimlerde kriz ve afet yonetim calismalarinin mevzuat acisindan degerlendirilmes, *Selcuk Universitesi Sosyal ve Teknik Arastirmalar Dergisi* (12): 195–219.

Cetina, T. (2014). The prediction of the critical factor of safety of homogeneous finite slopes subjected to earthquake forces using neural networks and multiple regressions, *Geomech Eng* **6**(1): 1–15.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation, *arXiv preprint arXiv:1406.1078* .

Cilli, M. (2007). Insan hareketlerinin modellenmesi ve benzesiminde temel bilesenler analizi yonteminin kullanilmasi, *Yayinlanmamis Doktora Tezi. Hacettepe Üniversitesi Saglik Bilimleri Enstitusu, Ankara* .

Cui, S., Yin, Y., Wang, D., Li, Z. and Wang, Y. (2021). A stacking-based ensemble learning method for earthquake casualty prediction, *Applied Soft Computing* **101**: 107038.

Demirci, E. and Karatli, M. (2023). Kripto para fiyatlarinin lstm ve gru modelleri ile tahmini, *Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* **10**(1): 134–157.

DeVries, P. M., Viégas, F., Wattenberg, M. and Meade, B. J. (2018). Deep learning of aftershock patterns following large earthquakes, *Nature* **560**(7720): 632–634.

Donges, N. (2018). Recurrent neural networks and lstm, *Towards Data Science* .

Ergunay, O. (2007). Turkiye'nin afet profili, *TMMOB Afet Sempozyumu* **5**(7): 1–14.

Gavcar, E. and Metin, H. M. (2021). Hisse senedi degerlerinin makine ogrenimi derin ogrenme) ile tahmini, *Ekonomi ve Yonetim Arastirmalari Dergisi* **10**(2): 1–11.

Gruber, N. and Jockisch, A. (2020). Are gru cells more specific and lstm cells more sensitive in motive classification of text?, *Frontiers in artificial intelligence* **3**: 40.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory, *Neural computation* **9**(8): 1735–1780.

Holliday, J. R., Chen, C.-c., Tiampo, K. F., Rundle, J. B., Turcotte, D. L. and Donnellan, A. (2007). A relm earthquake forecast based on pattern informatics, *Seismological Research Letters* **78**(1): 87–93.

Iatan, I. (2015). Predicting the earthquake magnitude from seismicity indicators using a probabilistic neural network, *The 13th Workshop Of Scientific Communications*, Vol. 23, p. 79.

Isci, C. (2008). Deprem nedir ve nasil korunuz?, *Journal of Yasar University* **3**(9): 959–983.

Janardhanan, D. and Barrett, E. (2017). Cpu workload forecasting of machines in data centers using lstm recurrent neural networks and arima models, *2017 12th international conference for internet technology and secured transactions (ICITST)*, IEEE, pp. 55–60.

Jang, J.-S. (1993). Anfis: Adaptive-network-based fuzzy inference system, *IEEE transactions on systems, man, and cybernetics* **23**(3): 665–685.

Jozinovic, D., Lomax, A., Štajduhar, I. and Michelini, A. (2020). Rapid prediction of earthquake ground shaking intensity using raw waveform data and a convolutional neural network, *Geophysical Journal International* **222**(2): 1379–1389.

Kamath, R. and Kamat, R. (2017). Earthquake magnitude prediction for andaman-nicobar islands: Adaptive neuro-fuzzy modeling with fuzzy subtractive clustering approach, *J. Chem. Pharm. Sci* **10**(3): 1228–1233.

Karakus, U. (2014). Deprem yasamis ve yasamamis ogrencilerin deprem algilarinin metafor analizi ile incelenmesi, *Dogru Cografya Dergisi* **18**(29): 97–116.

Karci, M. and Sahin, I. (2022). Derin ogrenme yontemleri kullanılarak deprem tahmini gerceklestirilmesi, *Artificial Intelligence Studies* **5**(1): 23–34.

Ketin, I. (1994). *Genel jeoloji: Yerbilimlerine Giris*, ITU Vakfi Yayinlari.

Kostadinov, S. (2017). Understanding gru networks. towards data science, *Towards Data Science, Towards Data Science* **16**.

Malkocoglu, A. B. V., ORMAN, Z. and ŞAMLI, R. (2022). Comparison of different heuristics integrated with neural networks: A case study for earthquake damage estimation, *Acta Infologica* **6**(2): 265–281.

Mallouhy, R., Abou Jaoude, C., Guyeux, C. and Makhoul, A. (2019). Major earthquake event prediction using various machine learning algorithms, *2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*, IEEE, pp. 1–7.

Mirrashid, M., Givehchi, M., Miri, M. and Madandoust, R. (2016). Performance investigation of neuro-fuzzy system for earthquake prediction.

Mohankumar, K. and Sangeetha, K. (2018). A study on earthquake prediction using neural network algorithms, *International Journal of Computer Sciences and Engineering* **6**(10): 200–204.

Moustra, M., Avraamides, M. and Christodoulou, C. (2011). Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals, *Expert systems with applications* **38**(12): 15032–15039.

Narayanakumar, S. and Raja, K. (2016). A bp artificial neural network model for earthquake magnitude prediction in himalayas, india, *Circuits and Systems* **7**(11): 3456–3468.

Natarajan, Y., Wadhwa, G., Ranganathan, P. A. and Natarajan, K. (2023). Earthquake damage prediction and rapid assessment of building damage using machine learning., *2023 International Conference for Advancement in Technology (ICONAT)*, IEEE, pp. 1–5.

Nicolis, O., Plaza, F. and Salas, R. (2021). Prediction of intensity and location of seismic events using deep learning, *Spatial Statistics* **42**: 100442.

Niksarlioglu, S. and Kulahci, F. (2013). An artificial neural network model for earthquake prediction and relations between environmental parameters and earthquakes, *International Journal of Geological and Environmental Engineering* **7**(2): 87–90.

Ozturk, S. and Alkan, H. (2022). Van golu ve civari (turkiye) icin deprem tahmini ve tehlikesi uzerine istatistiksel bir analiz ve degerlendirme, *Turk Deprem Arastima Dergisi* **4**(2): 191–209.

Pandit, A. and Biswal, K. C. (2019). Prediction of earthquake magnitude using adaptive neuro fuzzy inference system, *Earth Science Informatics* **12**: 513–524.

Phi, M. (2018). Illustrated guide to lstm's and gru's: A step by step explanation, sep 24, 2018, *URL: https://towardsdatascience. com/illustrated-guide-tolstms-and-gru-sa-step-by-step-explanation-44e9eb85bf21 (visited on 11/19/2020)* .

Reyes, J., Morales-Esteban, A. and Martinez-Alvares, F. (2013). Neural networks to predict earthquakes in chile, *Applied Soft Computing* **13**(2): 1314–1328.

Sahingoz, O. K., Çebi, C. B., Bulut, F. S., Fırat, H. and Karatas, G. (2020). Deep learning based security management of information systems: A comparative study.

Sebatli-Saglam, A. and Cavdur, F. (2022). Earthquake intensity estimation via an artificial neural network: Examination of different network designs and training algorithms, *Journal of the Faculty of Engineering and Architecture of Gazi University* **37**(4): 2133–2145.

Sen, S. (2023). Kahramanmaras depremlerinin ekonomiye etkisi, *Diplomasi ve Strateji Dergisi* **4**(1): 1–55.

Shcherbakov, M. V., Brebels, A., Shcherbakova, N. L., Tyukov, A. P., Janovsky, T. A., Kamaev, V. A. et al. (2013). A survey of forecast error measures, *World applied sciences journal* **24**(24): 171–176.

Sheng, J., Mu, D., Zhang, H. and Lv, H. (2015). Seismotectonics considered artificial neural network earthquake prediction in northeast seismic region of china, *The Open Civil Engineering Journal* **9**(1).

Shibli, M. (2011). A novel approach to predict earthquakes using adaptive neural fuzzy inference system and conservation of energy-angular momentum, *International Journal of Computer Information Systems and Industrial Management Applications. ISSN* pp. 2150–7988.

Sivaiah Bellamkonda, L., Vedantham, R. and Vemula, M. K. (2021). An enhanced earthquake prediction model using long short-term memory, *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* **12**(14): 2397–2403.

Utkucu, M., Kurnaz, T. F. and Ince, Y. (2023). The seismicity assessment and probabilistic seismic hazard analysis of the plateau containing large dams around the east anatolian fault zone, eastern turkiye, *Environmental Earth Sciences* **82**(15): 371.

Vasti, M. and Dev, A. (2020). Classification and analysis of real-world earthquake data using various machine learning algorithms, *Advances in Data Sciences, Security and Applications: Proceedings of ICDSSA 2019*, Springer, pp. 1–14.

Wang, Q., Guo, Y., Yu, L. and Li, P. (2017). Earthquake prediction based on spatio-temporal data mining: an lstm network approach, *IEEE Transactions on Emerging Topics in Computing* **8**(1): 148–158.

Wang, X., Zhong, Z., Yao, Y., Li, Z., Zhou, S., Jiang, C. and Jia, K. (2023). Small earthquakes can help predict large earthquakes: A machine learning perspective, *Applied Sciences* **13**(11): 6424.

Wang, Y., Chen, Y. and Zhang, J. (2009). The application of rbf neural network in earthquake prediction, *2009 Third International Conference on Genetic and Evolutionary Computing*, IEEE, pp. 465–468.

Wang, Z., Zentner, I., Pedroni, N. and Zio, E. (2017). Adaptive artificial neural networks for seismic fragility analysis, *2017 2nd international conference on system reliability and safety (ICSRS)*, IEEE, pp. 414–420.

Weiss, G., Goldberg, Y. and Yahav, E. (2018). On the practical computational power of finite precision rnns for language recognition, *arXiv preprint arXiv:1805.04908* .

Wen, W., Ji, D., Zhai, C. and Dong, K. (n.d.). Prediction on the amplitudes of seismic underground motions based on deep neural network.

Wiemer, S. and Wyss, M. (1994). Seismic qiescence before the landers (m= 7.5) and big bear (m= 6.5) 1992 earthquakes, *Bulletin of the Seismological Society of America* **84**(3): 900–916.

Yildirim, O. (2010). Analysis of earthquakes by means of data mining methods.

Yildirim, O. (2018). A novel wavelet sequence based on deep bidirectional lstm network model for ecg signal classification, *Computers in biology and medicine* **96**: 189–202.

Zakeri, N. S. S. and Pashazadeh, S. (2015). Data mining techniques on earthquake data: Recent data mining approaches, *Improving Knowledge Discovery through the Integration of Data Mining Techniques*, IGI Global, pp. 183–199.

Zhou, F. and Zhu, X. (2014). Earthquake prediction based on lm-bp neural network, *Proceedings of the 9th International Symposium on Linear Drives for Industry Applications, Volume 1*, Springer Berlin Heidelberg, pp. 13–20.