Republic of Iraq
Ministry of Higher Education and Scientific Research
University of Babylon
College of Science for Women
Department of Computer Sciences

A proposed Approach to Predict the Forest Fires

A Thesis

Submitted to the Council of the College of Science for Women at the University of Babylon in Partial Fulfillment of the Requirements for the Degree of Master in Science \ Computer Science

By
Nawras Yahya AL-KHafaji

Supervised By

Prof. Dr. Samaher Al-Janabi

2023 A.D. **1444 A.H**

بســم الله الرحمـن الرحيم

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ۞ خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ ۞ اقْرَأْ وَرَبُّكَ الْأَكْرَمُ ۞الَّذِي عَلَّمَ بِالْقَلَمِ ۞ عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمْ۞

صدق الله العلي العظيم

سورة العلق: الاية [1- 5]

Dedication

To God, my Lord, my Creator, my Highest, and my Hope.

To the Messenger who reached the valley of safety and advised the nation.

To the Prophet of Mercy and Light of the Worlds "Our Master Muhammad, may God's prayers and peace be upon him, and his God."

To my father ***(Yahya Hussein Al-Khafaji)*** *and mother* ***(Fawzeya Shaker Ai-Shemary)*** *who were the reason for my success by praying for me*

To my dear husband who supported me a lot during the study period habiby ***Ali Hassan Al-KHafaji****.*

My children have the blessing of my eyes, whom I fell short throughout my studies, and they are ***Suhaila, Ali Al-Akbar, Fatima, Zainab, and Dorar.***

To all my family and friends who have spared no effort in encouraging and supporting me.

For the College of Science for Girls, especially the Computer Department, this gave me the opportunity to complete my master and the faculty who were the reason for giving me knowledge.

Nawras Yahya AL-KHafaji
2023

Acknowledgment

First of all, we thank **God** for the most mercy for enabling us to present this project in the best form that we wanted it to be.

Basically, I would like to thank the greatest woman, my supervisor **Prof. Dr. Samaher Hussien Al-janabi**, for her helpful, guidance, supervision, and untiring efforts during the course of this work.

Thanks, and Gratitude to all my professors and all the staff of the **Department of Computer Sciences \ College of Science for Women \ University of Babylon** for their help.

Nawras Yahya AL-KHafaji
2023

Supervisors Certification

We certify that this thesis entitled ***"A proposed Approach to Predict the Forest Fires"*** *was done by (* ***Nawras Yahya Al-Khafaji*** *) under my supervision.*

Signature:
Name: Pro. Dr. Samaher Al - Janabi
Date: / / 2023
Address: University of Babylon / College of Science for women

The Head of the Department Certification

In view of the available recommendations, I forward the dissertation entitled ***" Aproposed approach to Predict the Forest Fires"*** for debate by the examination committee.

Signature:
Name: Assist Prof Dr. Saif Al-Alaq
Date: / / 2023
Address: University of Babylon / College of Science for Women

Abstract

Forests are considered one of the main wealth on the surface of the globe and have a fundamental role in that they provide a suitable environment for the life of many living organisms in addition to maintaining their ability to have a moderate climate on the surface of the globe, on the other hand, forest fires are one of the most important disasters that The Earth is exposed to it from time to time due to several phenomena, including natural and industrial ones, which may be due to the intervention of humans or technology.

This thesis attempts to build a model (Predictor Knowledge Constrains Gated Recurrent Unit (PKC-GRU)) based on the formulation of mathematical models for the basic determinants that may be the cause of fires in forests, where a named model was built (PKC-GRU), which consisted of four and consisted of four stages: The first stage included various databases for scientific research on forest fires , which included many characteristics, and it was necessary to determine which of these characteristics is more important in the occurrence of these fires in those forests. As for the second stage, it is considered the basis of the proposed model, through which the most important characteristics were used to build five determinants of varying importance and were in the form of mathematical formulas based on the most important

characteristics. The third stage of building the predictor received from training and testing, the last stage of the proposed model determining the quality of the data resulting from this treatment using evaluation scales. The results were encouraging and acceptable, compared to the rest of the networks, where it trains faster and performs better on less training data and thus is easier to train, since (GRU) contains only two gates: the update gate and the reset gate just like the gates of (long short term memory (LSTM)) where these gates are trained in (GRU) to selectively filter out any irrelevant information while keeping what is useful. These gates are basically vectors that contain values from 0 to which will be multiplied by the input data and/or hidden state. A value of 0 in the gate vector indicates that the corresponding data in the input or hidden state is not significant and therefore returns as zero On the other hand, a value of 1 in the gate vector means that the corresponding data is important and will be used and when we compared the approved database with the database of Algeria in its two regions, Bejaia and Sidi bil Abess, it showed that the Algeria database has a weak response to the determinants of knowledge, because the description of the approved database is regression, while the Algeria database is classification and regression.

Table of Contents

List of Tables

List of Figures

List of Algorithms

List of Abbreviation	
BiLSTM	Binary long short term memory
CMT	Influence Temperature on Forest Fire Occurrence
CRH	The Role Relative Humidity in Forest Fires
CTR	The Impact of Diurnal Temperature Variation on Forest Fire
CWS	The Influence of Wind Speed of Forest Fires
FF	Forest Fires
FFMC	Fine Fuel Moisture Code
FWI	Fire Weather Index
GA	Genetic Algorithm
GIS	Geographic Information System
GNN	Graph neural network
GRU	Gated Recurrent Unit
ISI	Initial Spread Index
JNN	Jacobian Neural Network
KC	Knowledge constrains
KG	knowledge graph
LSTM	Long short term memory
MAE	mean absolute error
MARS	Multivariate adaptive regression splines
RBF	Radial basis function
RH	Relative humidity
RMSE	Root mean square error
ROC	Receiver Operating Characteristics
SVM	Support vector Machine

TCMs	Temporal Convolutional Machines
$Temp^-$	Highest temperature - temperature for each features
WOA	Whale Optimization Algorithm
X	x-axis spatial coordinate
Y	y-axis spatial coordinate

Chapter One:

General Introduction

Chapter One: General Introduction

1.1 Introduction

Forests are the basis for some of the densest and most diverse ecosystems for life on land, and they cover about 4.06 billion hectares or 31% of the Earth's surface. Therefore, forest fires are as old as forests, so the ecosystem, biodiversity, and forest wealth in the whole region are seriously threatened by the impact of forests. In addition, during the summer, forests are covered with pricks and thorny leaves. When it does not rain for months, fires will ignite at the slightest spark.

Forest fires consist of natural causes in addition to human causes: ***(a) Natural Causes*** - many forest fires start as lightning that ignites tree fires, however, other weather conditions in terms of high temperatures and dryness (low humidity) provide favourable conditions for starting a fire, but rain extinguishes such fires without causing serious damage**, (b)** ***Human-made causes*** - fire caused by human activity occurs when an ignition source, such as an electric spark or other source of ignition, comes into contact with combustible material. In general, there are three basic categories that forest fires can be divided into **(a)** Natural or out-of-control forest fires **(b)** fires from heat generated in burning trash and other biomes during the summer as a result of human error (human neglect); and **(c)** forest fires started intentionally by a local resident, Forest fire (Caton-Kerr et.al.,2019), The main benefits of forests can be summarized as air purification because trees absorb odors and toxic gases such as nitrogen oxide, ammonia, sulfur dioxide, and ozone by trapping them in their leaves and bark. It also traps suspended particles in the air according to (Gnusov et al., 2020), and it was found that forest density has a significant impact on the duration and spread of forest fires in the environment, while wind speed has a significant impact only on the duration of the fire, which indicates that Human activity has an impact on the spread of fires, (Rasooli, et al.,2021),(Abdi, et al.,2018).

Deep learning (Janiesch et,al.,2021) The most intelligent systems that use machine learning to provide artificial intelligence capabilities. Machine learning refers to a system's ability to learn from training data for a given problem in order to automate the process of creating an analytic model and completing related activities. Artificial neural networks serve as the foundation for the machine learning concept known as "deep learning", a subset of artificial intelligence in which algorithms and programs are tested by a computer using self-improvement and self-development (Das and Roy 2019), and deep learning is a type of machine learning that uses recombination Hierarchical features to extract relevant data and then learn the patterns contained in the data.

Neurocomputing, (Yang & Shami,2020) algorithms for machine learning have been widely used in a variety of applications and fields to meet various tasks, a machine learning model's hyper-parameters must be adjusted, the performance of machine learning models is directly impacted by selecting the right hyper-parameter configuration. It typically calls for a solid grasp of machine learning techniques as well as appropriate hyper-parameter optimization approaches.

The prediction could be defined as the task of data analysis to predict unknown values of the prediction target feature. It includes a classification task for class label prediction and a numerical prediction where the task is to predict continuous values or ordered values. Type of target attribute specifies if the problem is classified with binary values or numerical prediction with continuous values. Many statistical methodologies are used for numerical prediction and regression analysis is most often used (Basavaraju et al. 2019).

Finally, forest fire is one of the most important challenges facing the world today as a result of the development of technology. Where it can be defined from several aspects in terms. This research deals with intelligent predictive design to address this phenomenon.

1.2 Problem Statement

The main problem of this thesis is forest fire prediction based on a set of determinants and prediction techniques of type GRU, the main problem of this work is predicting the area. One of the Neurocomputing prediction techniques is gated recurrent unit(GRU) characterized by many features that make it the best. These features (i.e., GRU give high accuracy results and work with data / stream on data in real-time also it contains memory therefore, we will use it in this thesis.

1.2.1 Research Questions

- How knowledge Constraints can be useful in building Predicator?
- How can we combine two technologies Neurocomputing and knowledge Constrains to create a multi-layer model?
- What is the beneficial result from building predictor by a combination between KC and GRU?

1.3 Thesis Objectives

- Determine the Main Rules (i.e., Knowledge Constrains) for environments and Neurocomputing.
- Design Knowledge Constrains Gated Recurrent Unit (KC_GRU) for predicting the area of forest fires.

1.4 Literature Survey

The problem of forest fires prediction is one of the key issues related directly to people's lives and the continued healthful life generally (Hooker,S. A., Masters, K. S., & Park, C. L. 2018). In this section of the thesis, we will attempt to review the works of previous researchers in the same area of our issue and compare works with seven basis points, (i.e “authors (s), name of database/dataset used in that article, preprocessing method, methodology suggest solving that problem, main measures used to evaluate results, advantages of that methodology and disadvantages”.

(Lai, C., et ., al 2022), propose a forest fire prediction method that uses a sparse autoencoder-based deep neural network and a novel data balancing procedure dataset obtained from the Montesinho Natural Park of Portugal, could predict large-scale forest fires more accurately, and reduces the mean absolute error by 3-19.3 and root mean squared error by 0.95-19.3 prediction performance could be improved if additional information and more data are available, this work is similar to my work us in evaluation measures and deep learning but different with my work in data set using to predicte to forest fires.

(Gayathri, S., Karthi, P. A., & Sunil, S. 2022), Suggested a method using image processing unit and grayscale to detect forest fires using CNN deep learning algorithms and after detecting the fire an alert is sent to control the forest team along with the site. They also integrated Google's Firebase to send alerts through mobile phone or IoT device notifications and mainly engaged in regards to a brief introduction on the Backwoods fire was related work on strategies and different frameworks in the Timberland fires and a conversation on computational reasoning and AI computations followed by an audit of projected frameworks The site of using measure metrics such as accuracy, F1 score, the accuracy score of the train data is 96% and the accuracy score of the valid or the test data is 92%.

(Khan, S., and Khan, A. 2022), proposed a deep learning method called FFireNet, by taking advantage of the MobileNetV2 pre-trained convolutional base and adding fully connected layers to solve the new task, which is the forest fire recognition problem. The performance of the proposed fire and no-fire classification solution was evaluated using different performance metrics and compared with other CNN models, the results show that it has an accuracy of 98.42%, an error rate of 1.58%, a recovery of 99.47%, and an accuracy of 97.42% in the classification of fire images and non-fire images. The results of the proposed approach are promising regarding the problem of forest fire classification given Unique forest fires Detection Dataset Performance evaluation on wildfire dataset and comparison with NASNetMobile, Xception InceptionV3, and ResNet152V2.

(Kukuk, S.B., & Kilimci, Z.H. 2021) presented a comprehensive examination of object identification methods, deep and hybrid deep learning models, and traditional machine learning algorithms for wildfire detection and used evaluation metrics that are use accuracy, precision, recall, F-measurement, and intercept over union (IOU). This work is similar to mine in terms of forest fire prevention, but it differs from my work by using one of the convolutional neural network algorithms and the results of the experiment show that convolutional neural networks outperform other methods with an accuracy of up to 99.32%.

(Elvan et.al.,2021) proposed based on the "Forest Fires and the Law" guide from the Food and Agriculture Organization (FAO), which was founded by lawyers and covers the fundamental topics of definitions, institutional setup, inter-institutional coordination planning, monitoring, and assessment, prevention and preparedness, detection and early warning and suppression, participatory and community-based approaches, and institutional setup, monitoring, and assessment In terms of prediction, this work and research are comparable, but this work makes use of a fire detection system and a multipurpose artificial intelligence framework., whereas the study uses evaluation measures and different data for prediction, and this work uses a fire detection system.

(Singh et. al.,2021), predicated forest fires depending small data set using three options for “Cascade Correlation Network” (CCN), “Radial Basis Function” (RBF) and “Support Vector Machine” (SVM) the preprocessing depending “SPARK” and “PySpark” were utilized to carry out the prediction process' data segmentation and feature selection.

(Li, X et al., 2021) designed network based three-type models (LSTM) to predict the rate of fire propagation, and to explore the interaction between fire and wind. In order to train these LSTM based models and validate their prediction effectiveness. There are three progressive models based on LSTM created, which are called CSG-LSTM, MDG-LSTM, and FNU-LSTM. The feature of FNU-LSTM is further demonstrated. By performing comparison experiments with normal LSTM and other LSTM-based models, the result was that CSG-LSTM takes about 100 iterations and 13 minutes to reach. MDG-LSTM model, it takes about 100 iterations and takes 160 minutes to reach the maximum value of convergence for the speed of fire propagation, FNU-LSTM model, it takes about 10 iterations and 20 minutes to reach the value of convergence for the speed of fire propagation, Entropy Loss equation.

(Benzekri, W et.al.,2020) proposed system based on collecting environmental wireless sensor network data from Forest and forest fire prediction using artificial intelligence, specifically deep learning (DL) A combination of such a system based on the concept of an Internet of Things (IoT) consisting of a low power wide area network (LPWAN), fixed or mobile sensors and a good suitable model for deep learning gave the GRU model 99.89% accuracy and a loss function value of 0.0088. It is workable Similar using the same algorithm but different evaluation metrics.

(Al-Kahlout et. al.,2020), used automated tools to use local sensors such as those provided by weather stations, and many fire indexes, such as the Forest Fire Weather Index, have been developed to account for this (FWI), It has a high accuracy of, as well as lower data collection costs and better fire resource management, thanks to this data. To predict the burned area of a forest fire, only a neural network (JNN) method has been proposed. This research is similar to previous Forest Fire Database work, but instead of using traditional prediction metrics, it employs a neural network approach.

(Wang, G., et .,al 2019),Improving the effectiveness and accuracy of early forest fire detection, conducting a self-collection of forest fire data and two real forest fire observation videos Five forest fire monitoring video sequences collected from the webpage were used, the method of accurately dividing the flame area at the pixel level in the forest Fire results show its ability I have to work in various conditions of fire and difficult lighting.

This work is similar to my work in terms of its use of artificial intelligence and deep learning techniques, but differs from my work in terms of pre-processing and evaluation measures for the results.

(Firoz and Laxmi 2018), proposed a method of determination to forest fires using determinants such as shooting frequency. It was broadcast on five networks (with thighs frequency). Monthly average analysis, Rainfall, The speed of the wind, in addition to evaporation, the severity of the climate/weather, Examine the Pearson correlation coefficient, as their work is similar in terms of a data type but differs in its evaluation of climatic anomalies, Where we took a very long time compared to a number of researchers, and we did not choose the algorithm randomly, but we compared it to many researchers in the scientific field.And we will conclude through comparison the work that is close to our work in terms of the database used as well as the proposed algorithm and find weaknesses and strengths and compare with them.

Table 1 shows a comparison between the previous works of numbers of researchers in the same field and consists of seven points that include the name of the authors, the name of the database/dataset used in thesis, the pre-processing method, the methodology proposed to solve this problem, the main metrics used to evaluate the results, advantages and disadvantages of that methodology.

Table 1.1: Compare among the Previous Works

Name ofauthors	Dataset\Database	Preprocessing	Methodology	Evaluation measures	Advantage	Disadvantage
(Lai, C., et ., al 2022)	▪ dataset collected from the Montesinho Natural Park of Portugal. ▪ dataset is seriously imbalanced	▪ sparse autoencoder-based deep neural network ▪ selection of features ▪ imbalanced data	▪ Benchmark Data for Forest Fire Prediction ▪ Sparse Autoencoder ▪ Deep Neural Networks	▪ mean absolute error by 3–19.3 and root mean squared error by 0.95–19.3	▪ better benefit the management of wildland fires in advance and the prevention of serious fire accidents.	▪ Selection of features that are most pertinent to the prediction task is a prerequisite, and a distribution of data that is severely unbalanced
Gayathri, S., Karthi, P. A., & Sunil, S.	▪ • input image pixel-level accuracy	▪ image processing module, grey scalability, and mathematical climate	▪ CNN BiLSTM	▪ ccuracy,, F1 score	▪ rapid action to put out the fire; spatial feature detection employing forward and reverse order	▪ Various geographic circumstances
Khan, S., and Khan, A. (2022),	▪ designating pictures ,950 are part Fire instance and the remaining	▪ localization ▪ classification ▪	▪ CNN-based FFireNet	▪ Accuracy ▪ ER (minimum error rate.),TP,(true positive,)FT,F,	▪ • The forest fire detection dataset, which resolves the classification issue of differentiating between photos with and without fire.	▪ unique forest fire detection dataset.

Singh et .al.,(2021)	Elvan et. al .,(2021)	Kukuk, S.B., & Kilimci, Z.H.(2021)
• the datad set The prediction process' data segmentation and feature selection were carried out using SPARK and	• Using a reference from the Food and Agriculture Organization (FAO), "Forest Fires and the Law," a dataset in Turkey pertaining to Turkish forest fire legislation was	▪ Google image search is used to create the data collection. ▪ Data sets received by the webcam are identified as video picture data sets.
• decreased effectiveness in huge datasets as a result of model overfitting issues	• lawyers and addresses the essential definitional concerns	▪ Image analysis based on fire detection ▪ methods for classifying objects and identifying objects
• CCN RBF SVM	• • rehabilitation, interinstitutional coordination, planning, monitoring, and evaluation; prevention and preparedness	▪ SVM,CNN,RF,CNN-GRU,(CNN-LSTM,Faster Recurrent-Convolutional Neural Network (Faster R-CNN),Single Shot Detector (SSD),
• SVM	• The primary reference point being the Turkish Constitution	▪ Accuracy (AC), f-measure (FM), precision ▪ (PR), mean average precision
• accurate predictions	• This study's objective was to provide an overview of Turkish forestry law and practices as they currently stand in relation to FAO recommendations.	▪ usage in open spaces, scale of the hazards, and installation costs that are less expensivefurther exploitation.
• The frequency and severity of forest fires are influenced by a region's climatic features.	• the results demonstrate of how stringent the regulations, a lack of administrative safeguards and a lack of public knowledge	▪ n the field of object detection, this task is extremely difficult. ▪ identify the fire when interior fire detection devices are ineffective

Al-Kahlout et. al .,(2020)	W Benzekri et.al.,(2020)	Li, X et al., (2021)
▪ Dataset using sensor and FWI	• using deep learning to get data from forest wireless sensor networks	▪ The information is gathered using an infrared camera mounted on a drone and simultaneously recorded using an anemometer.
▪ Quick detection is essential for managing this problem.	• • The Low Power Wide Area Network (LPWAN), which is the foundation of the Internet of Things (IoT).	▪ Configuration: Calculating Fire Spreading Rate from Infrared Image Sequences
▪ ▪ FWI(fire weather index) ▪ JNN(just neural network)	• Artificial intelligence, the LoRa IoT system, and the Fire Weather Index are all used in deep learning.	▪ CSG-LSTM ▪ MDG-LSTM ▪ FNU-LSTM
▪ ANN(artifitial neural network) ▪	• binary cross-entropy (BCE)	▪ Entropy Loss equation
▪ High Resolution 98.75 Enhanced management of firefighting resources low expenses for gathering data	• • Possibility of a self-sufficient, real-time environmental monitoring system for elements that affect the danger of forest fires	▪ Quality of the model training, and Prediction Accuracy
▪ Big data that is inefficient increases computation time.	• high false alarm rate brought on by external factors such as adjacent fog, dust, shadows, etc.	▪ Measurement of the external environment's wind speed is challenging.

Firoz and and Laxmi (2018)	(Wang, G., et ..al 2019),
▪ Dataset from(five grids) using Remote Sensing and GIS in India	▪ five forest fire monitoring video sequences collected from the Web pages.
▪ Deciduous Broadleaf Forest	▪ an accurate flame zone segmentation approach at the pixel level
▪ Remote Sensing and GIS	▪ Deep Learning(CNN), Artificial Intelligence,
▪ Analyze forecasted climate ▪ anomalies	▪ true positive (TP) ▪ false positive (FP)
▪ Forecasted climate anomalies should be examined volume up content copy share star border	▪ Using image classification to perform dense prediction tasks like fire region segmentation might increase the effectiveness and precision of spotting early forest fires.
▪ This issue was represented as a regression job.	▪ not detect early forest fire

1.5 Thesis Layout

The remained chapters of this thesis organization as the following:

- *Chapter two:* gives an overview of the main theoretical background in this field.
- *Chapter three:* explains a build of the predictors based on knowledge constrains.
- *Chapter four:* illustrates the implementation of the KC-GRU model and the results of the cases study.
- *Chapter five:* shows conclusions of this work together with some recommendations for work in this area in the future.

Chapter Two:

Theoretical Background

Chapter Two: Theoretical Background

2.1 Introduction

This chapter will present the main theoretical scale grouped related to the problem that was introduced in Chapter One, which includes the main role of knowledge constriction in problem-solving, as well as an analysis of prediction algorithms and the main presenter of it from two perspectives. techniques for data mining the main sensor used to capture data in real-time related to forest fire will also be explained, as well as the evaluation measure used in this thesis.

2.2 Prediction Techniques

Prediction techniques can uncover by looking at the values of other variables, you can figure out what the unknown values of a target variable are. Furthermore, by extending the creation of predictions regarding the state of the future, and possibilities, prediction utilized construct future goals. overtime utilizing specific methods. In a nutshell, it forecasts future activity while taking into account all potential influencing factors. Prediction techniques are widely used in a variety of fields, including marketing, finance, telecommunications, healthcare, and medical diagnosis, to aid in the optimization of future decision-making, there are several types of prediction techniques, and this thesis will focus on the GRU.

2.2.1 Gated Recurrent Unit(GRU)

The GRU is the newer generation of Recurrent Neural networks, and it's the modified version of LSTM but with less complexity. GRUs use the hidden state to transfer information instead of cell state. It also only has two gates, a reset gate and update gate.

Update Gate: The update gate selects information that needs to be added and / or ignored. It's similar to the LSTM input gate.

Reset Gate: The reset gate is used to decide how much past information to forget.Since GRUs are smaller in operation, it's faster than LSTM networks. And the reason both LSTMs and GRUs are successful, is because their gating mechanism:Preserves contextual information and long-term sequences, Avoids gradient issues. The gating functions allow the network to modulate how much the gradient vanishes, and since it's being copied four times, it takes different values at each time step. The values that they take on are learned functions of the current input and hidden state. The algorithm GRU is considered one of the types of Nero computing, and it is a deep learning that consists of several layers that lead to high accuracy and is characterized by the presence of a number of weights for each layer, as the algorithm consists of three layers, which are the input and a number of hidden layers and the output. Accounts are also reduced,Traditional neural network models are susceptible to structural issues such gradient explosion and over-fitting, whereas the deep GRU neural network model has poor information processing capabilities over several hidden layers and low update efficiency (Wang et.al.,2019),Recurrent Neural Networks are a sub-branch of NNs that include typical RNNs, LSTMs, and GRUs as their core algorithms. Recurrent Neural Networks are a sub-branch of NNs that include typical RNNs, LSTMs, and GRUs as their core algorithms. Recall that the conventional RNN structure consists of input, hidden, and output layers, A popular sort of RNN is the GRU (Gated Repetition Unit) technique, which is used to process time series data and learn intricate dynamic temporal characteristics from input sequences.

A typical method for handling sequential data is long-term memory with backpropagation gradients through internal gates. Although it uses less processing power and performs as well as LSTM, GRU also exhibits strong performance. As a result, GRU excels in time series-related problems , Tasks involving time-series exhibit good performance for GRU(H Luo et.al.,2021)

RNNs have recurrent units in their hidden layer, which, in contrast to Feed Forward Neural Networks, enables the algorithm to analyze sequence data. This is accomplished by continuously mixing inputs from the current timestep with concealed states from past timesteps. RNNs have recurrent units in their hidden layer, which, in contrast to Feed Forward Neural Networks, enables the algorithm to analyze sequence data. This is accomplished by continuously mixing inputs from the current timestep with concealed states from past timesteps.

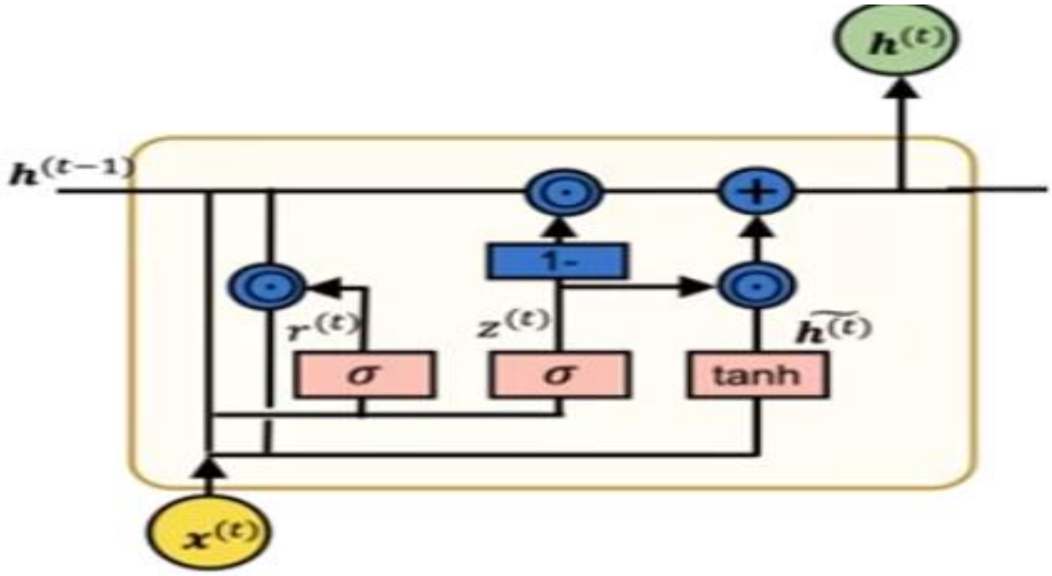


Fig (2.1): sructure of GRU This portal basically filters the most important , Gao, S. Et .al (2020)

information from among the important filters to reach the hidden layer

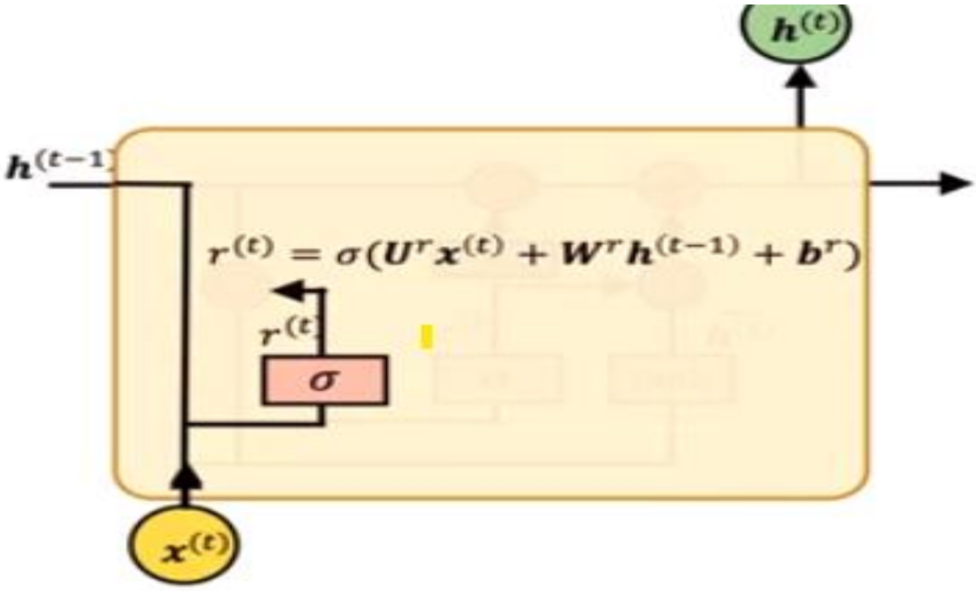
RNNs have recurrent units in their hidden layer, which in contrast to Feed Forward Neural Networks, enables the algorithm to analyze sequence data. This is accomplished by continuously mixing inputs from the current timestep with concealed states from past timesteps.

LSTM and GRU are comparable, but GRU has fewer gates. There is no separate cell state and just a hidden state is used for memory transfer across recurrent units. It also has three gates(Wang1 et.al.,2019),(Sajjad et.al.,2020):

- **Reset gate** :- This gate This portal basically filters the most important information from among the important filters to reach the hidden layer basically filters the most important information from among the important filters to reach the hidden layerThis portal basically filters the most important information from among the important filters to reach the hidden layerhelp to decide how to add the new information to the memory i.e how much past information if can forget ,it from the current input (x t) and the previous hidden state (h t-1) are multiplied by their respective weights, applied bias, and then transmitted through a reset gate. Step one determines which values should be discarded (0), remembered (1), or just partially maintained because the sigmoid function has a range between 0 and 1. (between 0 and 1). The previous hidden state is reset in step two by multiplying it by the results of step one, Gruber, N., and Jockisch, A. (2020).

$$r^{(t)} = \alpha(Ur \ X^{(t)} + W^{r} \ h^{(t-1)} + b^{r} \cdot \text{......} (2.1)$$

The Purpose of the Reset Gate,Reset gate's importance in determining how much information should be discarded is one of the main reasons for this. Given that the reset gate tends to classify irrelevant data before instructing the model to disregard it and continue without it, it would be reasonable to compare it to the forget gate in the LSTM, , Gao, S. Et .al (2020)

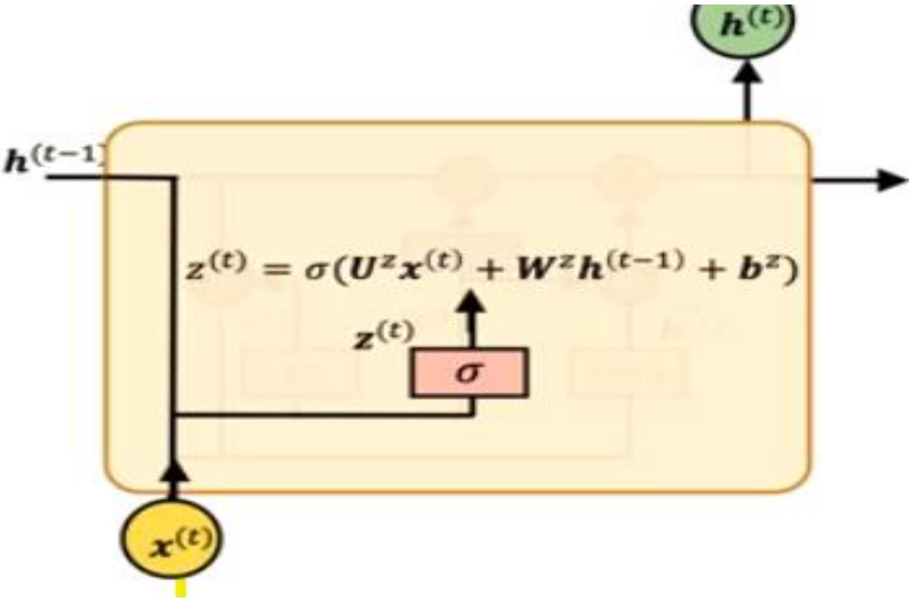


Fig(2.2): Structure of Reset Gate from GRU , , Gao, S. Et .al (2020)

- **Update gate** — help to decide what information from the previous time Step ht-1 can be taken forward to the next time step ht, weights and biases used to scale the vectors in this step produce a distinct sigmoid output. So after applying a sigmoid function to a combined vector, we subtract it from a vector that contains only 1s and multiply it by the prior concealed vector, uber, N., and Jockisch, A. (2020).

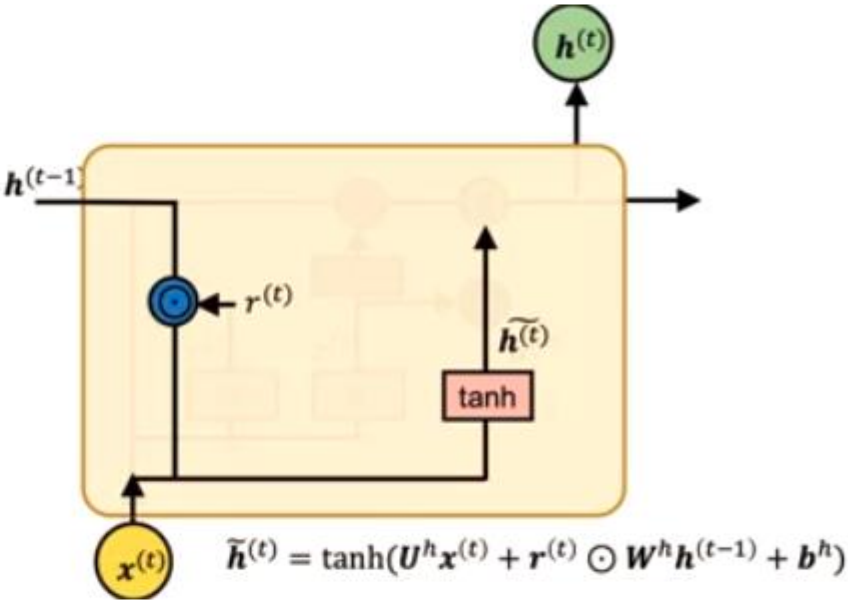
$$z^{(t)} = \alpha(U^z\ X^{(t)} + W^z\ h^{(t-1)} + b^z \quad \text{.... (2.2)}$$

The Purpose of the Update Gate, the update gate's primary job is to decide how much earlier information is best for future use. This function is crucial because the model may copy every single previous feature, which solves the fading gradient problem.



Fig(2.3): Structure of Update from GRU, Gao, S. Et .al (2020)

- **Hidden state candidate** — we just saw how update and reset gate work, but how do these help in updating the hidden state the outputs are mixed with new inputs (x t), multiplied by their respective weights, and added biases before passing through a tanh activation function. This is done after resetting a prior hidden state in step two, the new hidden state (h t) is created by multiplying the hidden state candidate by the output of an update gate and combining it with the previously changed hidden state (h t-1). Once the recurrent unit has processed the complete sequence, the process then repeats for time step t+1 and subsequent timesteps.



Fig(2.4):structure of hidden state from GRU, Gao, S. Et .al (2020)

Next, the process repeats for timestep t+1, etc., until the recurrent unit processes the entire sequence, Gruber, N., and Jockisch, A. (2020).

$h^{\sim(t)} = \tanh(U^h \; X^{(t)} + r^{(t)} \odot W^h \; h^{(t-1)} + b^h$ ..... (2.3)

$h^{(t)}=(1-z^{(t)}) \odot h(^{t-1)+}z^{(t)} \odot h^{\sim(t)}$ .... (2.4)

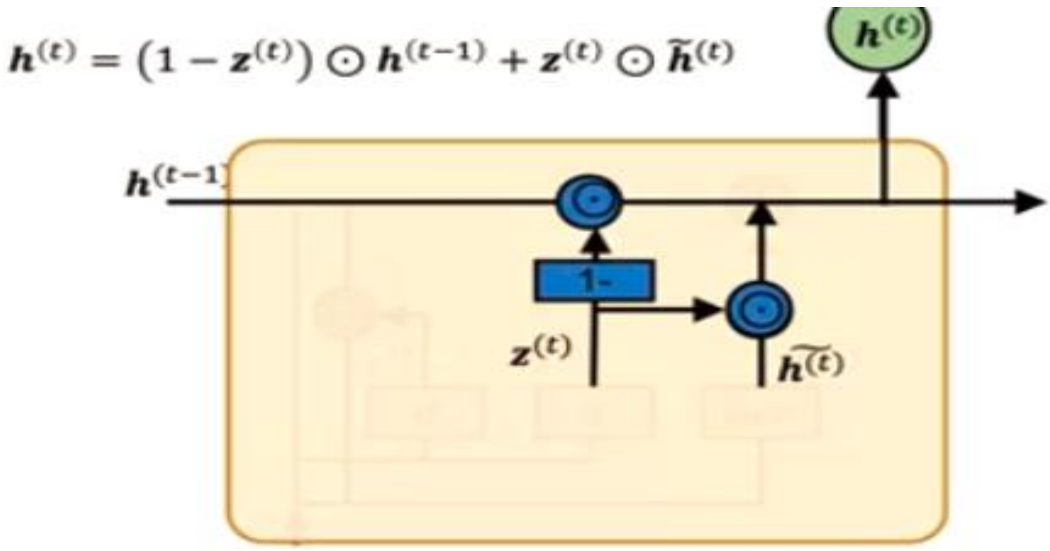


Fig (2.5): Structure of output of GRU, Gao, S. Et .al (2020

X(t): input vector, $h^{(t):}$ output vector, $h^{\sim(t):}$ candidate activation vector, z(t): update gate vector, r(t): reset gate vector, W, U, and b: parameter matrices and vectors, How to add new information? a new candidate state is created for holding the information, How to remove information from the hidden state? Know how update gate zt help to decide what information the previous time step t-1 can be taken forward to the next time step ht , multiplying ht-1 with zt give as releval information from the previous step Instead of having new information taking the complement of z^t i.e (1-z^t) and multiplying with $h^{\sim(t)}$

2.3 Multi Variances Analysis

2.3.1 Entropy

Entropy is a measure of purity that describes how uncertain, impure, or disordered a random variable is. It describes the impurity inside a random class . The measurement of impurity or unpredictability in the data points is called entropy. If all of the constituents in this situation fall into one class, it is said to be "Pure," and if not, the distribution is called "Impurity." It is calculated between 0 and 1, however depending on how many groups or classes are available in the data set, it may be greater than 1, still indicating a severe level of disorder. When all of the observations in a dataset belong to the same class, the dataset is devoid of impurity and randomness, and the entropy of that dataset is zero. This is the case when a dataset is made up of homogenous subsets of observations, Tsallis, C. (2022).

Clausius introduced the concept of "entropy" in 1865, and as a result, now referred to as "the principle of increasing entropy, Since Clausius only understood the idea of entropy in terms of "transformation" and "changing, system's total entropy value," he only mentioned the "phenomenon of growing entropy" (Wu and Wu (2020).

The unique aspect of Ludwig Boltzmann's work is that, in addition to providing a probabilistic approach to determining a system's entropy value, it also illuminates the concept of entropy's broader creative importance and worth through this calculation. Through his work, Boltzmann not only clarified the significance of entropy but also provided a probabilistic interpretation of its significance.Shannon's "Entropy of Information Sources." Clausius' principle of growing entropy was introduced more than 80 years after Shannon's "entropy information sources" theory, which is a related theory that is closely tied to the ideas of entropy and information.

In 1948, Boltzmann's statistical entropy theory directly inspired Shannon to develop his communication information theory. Boltzmann's statistical entropy theory provided Shannon with two tools: the statistical technique and the entropy equationa method for calculating the volume of information produced by the information source has been developed (Wu and Wu (2020).

Shannon's information entropy formula , Escolano, S. (2003)

$\mathrm{H} = -\sum_{i=1}^{n} pi \ \mathrm{Log2} \ Pi$ ....(2.5) where: $i = 1, 2, 3 \ldots n$; pi: probability of a event i $\Sigma pi = 1$, the Pi is has several possible meanings,It can be a measurement of the source's message's randomness, the source's message's a priori unpredictability, or the source's ability to generate messages (Ben-Naim, A. 2020).

Boltzmann made the brilliant connection between entropy and the following decade, while Shannon and Jaynes independently developed the connections with the theory of communications and, more generally, with the theory of information in 1948 and 1957. Since then, more than fifty novel entropic functionals have been published in the scientific and technological literature, (Tsallis, C. 2022), the letter confirmedin this message, entropy is the amount of information contained in each characteristic of a forest fire database to detect and predict a burned area.

2.3.2. Information Gain(IG)

The information gain formalism makes an effort to measure variation in the performance as a Gaussian distribution (Stephenson, M., et al ., 2020) , When conducting a process of gaining information, the idea of entropy is very important. However, since the information about the database is classified as having key qualities and necessary characteristics, the information that may be gained depends on the information in the database that is to be processed on. As it specifies the notion of Gain impurity information in the dataset and calculates the difference between the entropy before and after segmentation, it aims to uphold the concept of entropy, Tsallis, C. (2022), Nelson, J. D. (2005).

$$IG = \sum IG(target) - IG(all\ feature) \qquad \textbf{......(2.6)}$$

Information gain is not negative , When the split variable and target variable are switched, symmetric information gain ensures that the amount of information acquired does not change. Information gain determines how much the uncertainty is reduced after segmenting the dataset on a particular feature; as a result, the feature with the largest information gain is the most advantageous for classification.The ideal feature to be selected for splitting is the one with the maximum value of information gained,Information gain quantifies the impurity in class variables and compares the entropy before and after splitting.With symmetric information gain, the amount of information gained remains constant when the split variable and target variable are switched.

After segmenting the dataset on a particular feature, information gain determines how much the uncertainty is reduced; as a result, the feature that has the highest information gain is the most beneficial for classification.The ideal feature to be selected for splitting is the one with the maximum value of information gained,Information gain quantifies the impurity in class variables and compares the entropy before and after splitting.

2.4 Knowledge Constrains

The negative effects of global climate change are the result of increased droughts and hot summers in different parts of the region, as well as frequent forest fires.It can only be confirmed that humans frequently cause forest fires in various parts of the world, either directly or indirectly (for the purpose of, for example, expanding pastures or benefiting from insurance), Elvan et .al(2021).

Forest fires are one of the most dangerous environmental issues in the world because they spread over very large areas. This causes concern because such large fires may burn for days or even months and pose a number of risks to the environment, including the release of toxic gases like carbon monoxide, significant losses in plant and animal life, and harm to people and livestock. Because it plays a significant function as a regeneration, alteration, and air conditioning agent through transpiration and climate conditioning, including daily temperature, forest fires have the potential to positively influence ecosystems, Yet, the drawbacks of forest fires lay in their capacity to alter and ruin the tourist and aesthetic appeal of those locations, and the hazardous chemicals they produce go beyond a country's boundaries to neighbouring nations , Elvan et.al(2021).

The main natural cause of fires, in addition to the nearby that starts fires, is thought to be lightning. It is believed that this is due to the regular forset fires that occur in boreal woods and perhaps even in mountainous regions, particularly when coniferous trees are prominent and exposed. Nowadays, lightning is to blame for 30% of all fires in the nation. It is believed that the human factor (whether irresponsible or malicious) is to blame for at least 95% of fires. Human activities and infrastructure provide the energy needed to start fires. So, whether as a direct cause of ignition or in terms of fuel production, the human element has a considerable influence.

The main influence of meteorological factors on forest fires is seen above the analysis point of view; there are some factors that have an impact on each and every time a forest fire occurs, but the decision-making process does not depend solely on the causes of the fires; rather, it depends on a number of factors working in concert.

A number of factors, such as the "highest Temperature of the day," "Temperature range in diurnal," "Humidity the day's," "Wind Speed," "Precipitation," and other combinations, can have an impact on forest fires. Here are the five fundamental knowledge constraints equations explained, some of the data ***Yichun fire*** nearly a decade predict the forest fire meteorological data analysis, Yichun is the northern of Heilongjiang Province in China.Territory is hilly areas It is the temperate continental monsoon climate, annual rainfall 630 mm, with an average annual temperature of 1 degree, in January average temperature of minus 25 degrees, in July the average temperature of 21 degrees above zero, Dan Liu and Yanrong Zhang(2015).

- ***KC based on Temperature***

$$\boldsymbol{KC1} = \begin{cases} \left\{ \mathbf{1}/(\mathbf{1} + \left(\frac{\mathbf{1}}{\mathbf{5(20-}\boldsymbol{temp}\mathbf{)}}\right)^{\mathbf{4}} \right\} \; \boldsymbol{if\ Temp} < \mathbf{20} \\ \mathbf{1} \qquad\qquad\qquad\qquad \boldsymbol{otherwise} \end{cases} \qquad \textbf{(2.7)}$$

Where °C: An abbreviation for Celsius degree,temp: Temprature

KC1: temperature contribution

Temperature plays a major role in influencing the occurrence of fires a ccording to the temperature equation, the chance of fire outbreaks is zero if the temperature is low, and the rate of fire outbreaks increases if the temperature is greater or equal to 20. If (temp) is less than 20, the frequency of fires drops.KC based on “Temperature range in diurnal“ Variation on forest fire occurrence also In determining the contribution of the high value of fire risk (%), kc1 is determined by one factor. The highest temperature of the day is temp, with 14:00 having the greatest temperature. There is essentially minimal fire risk at 14:00 due to the low temperature, especially in the next 8 °C; There is a gradual increase in the fire threat and the temperature is between 8 and 12; Potential fires should receive extra attention between the hours of 12 and 20, due to the quick changes in the danger function during the low to high transition period; Above 20 °C, special attention will be paid, Dan Liu and Yanrong Zhang(2015)

- ***KC based on Diurnal Temperature***

$$\boldsymbol{KC2} = \begin{cases} \{\mathbf{1/(1 + (1/10(20 - (\boldsymbol{temp} - \boldsymbol{temp}^{-}))^{6}}\}\ \boldsymbol{if}\ (\boldsymbol{temp1} - \boldsymbol{temp}^{-}) < \mathbf{25} \\ \mathbf{1} \qquad \boldsymbol{otherwise} \end{cases} \quad \textbf{(2.8)}$$

Where °C: An abbreviation for Celsius degree, KC2=temperature Diurnal contribution

Using the fire risk equation (high daily temperatures), cloudy and foggy weather will make it more difficult for a fire to start; that is, if (temp-temp) < 25 overcast, rainy, and fog weather phenomena are more frequent; When the temperature ranges (temp-temp) > 25 high, the day is sunny, very hot, and the afternoon wind is very strong, there is also a very high likelihood of fires. The diurnal temperature high fire risk contribution value (%) only takes into account (temp-temp). The lowest temperature of the day is at 2:00, with (temp) and (temp) denoting the highest and lowest temperatures of the day, respectively. Under normal conditions, when (temp-temp^{-}) 12°C, cloudy, rainy, and foggy weather phenomena more, so difficult to fire; when (temp-temp¯) is between 12°C and 20°C, there is a significant increase in the degree of risk of fire; when (temp-temp¯) > 20°C, weather controlled by high pressure situation, the performance of sunny, daytime warming intense, afternoon the wind speed increases, the fire to maintain a higher state, Dan Liu and Yanrong Zhang(2015).

- ***KC based on Relative Humidity***

$$\boldsymbol{KC3} = \begin{cases} \{\mathbf{1/(1+(1/10(}\boldsymbol{RH}\mathbf{-20))^4}\}\ \boldsymbol{if}\quad \boldsymbol{RH} > \mathbf{15\%} \\ \mathbf{1} \qquad\qquad\qquad\qquad\qquad\qquad \boldsymbol{otherwise} \end{cases} \quad \textbf{..... (2.9)}$$

Where °C: An abbreviation for Celsius degrees,RH: relative humidity,

KC3:relative humidity contribution

relative humidity is less than 15%. from the application of the relative humidity equation, it was found that if the humidity is less than 15%, it indicates the occurrence of fires that reach their peak also A high fire danger number (%) is reflected in kc4, which is the relative humidity. RH is the 14:00 air relative humidity (percent). There is a low likelihood of a fire starting when the relative humidity is higher than 45%. The ratio of high fire risk starts to rise between 10% and 45%. Moreover, fires are most likely to start when the relative humidity is below 15%, Dan Liu and Yanrong Zhang(2015)

- ***KC based on Rain***

$$\boldsymbol{KC4} = \begin{cases} \{\mathbf{1/(1/(1+(}\boldsymbol{rain}\mathbf{)^3}\}\ \boldsymbol{if}\quad \boldsymbol{rain} > \mathbf{0} \\ \mathbf{1} \qquad\qquad\qquad\qquad\qquad\qquad \boldsymbol{otherwise} \end{cases} \quad (2.10)$$

where,°C: An abbreviation for Celsius degrees,KC4:rain contribution

More rainfall results in a decrease in the fire danger index, and vice versa, according to the application of the rain equation. The contribution value (%) for the 24 hours of high fire risk precipitation indicates kc5. Rain makes up the 24-hour precipitation factor. Changes in precipitation curve revealed a smooth declining trend as precipitation increased and the fire hazard index decreased. When there is less than 1mm of precipitation, you are in a high fire danger area, and the fire trend will not significantly slow down. The fire hazard would appear to be considerable once more if precipitation stopped now, Dan Liu and Yanrong Zhang(2012).

- ***KC based on Wind Speed***

$$\boldsymbol{KC5} = \begin{cases} \{\mathbf{1/(1/(1+1/12(7-\mathit{wind}))^{14}}\} & \boldsymbol{if \quad wind < 7} \\ \mathbf{1} & \boldsymbol{otherwise} \end{cases} \tag{2.11}$$

Where , wind:wind speed ,°C: An abbreviation for Celsius degrees,

KC5:wind speed contribution

since wind and rain commonly occur in the north at the same time, it is challenging to put out flames when the wind speed is over 7, and less likely to happen if it is below 7, kc5 is the only element that accounts for the percentage contribution of high fire danger. Around 14:00, the ground's 10m–12m height average wind speed is equal to wind. In the north, wind and precipitation typically happen at the same time, therefore wind speed measures are used to reduce the impact of its precipitation. Wind speeds of 3 m/s or less have minimal impact on the occurrence of forest fires; 3 m/s to 7 m/s cause the impact of forest fires to start growing quickly; and 8 m/s or more make it difficult to control after the fires have begun.

Chapter Three:

Building of Predicator Knowledge Constrains – GRU(PKC-GRU)

Chapter Three: Building of (PKC-GRU)

3.1 Introduction

This chapter builds a system called Knowledge Constrains forest fires with algorithm gated recurrent unit (PKC-GRU) to find the burned area . The main goal of this work is to find the new contra of forest fire then determined which one more effect in that problem. Therefore, we model called (KC-FF) is consists of stages: First stage; called Collecting dataset and Pre-processing , which include bank of data sets of forest fires, After that determined the features most important based on both (entropy and Information Gain) While the second stage include build knowledge constraints based on different features theses constraints are (KC1, KC2, KC3, KC4, and KC5) and Final stage related to evaluated the results of KC-FF model based on measures The value handling, and data set normalization; the second stage is the application of the GRU algorithm, which includes showing the input, hidden, and output layers with knowledge parameters that are either natural or industrial parameters, and the third stage is the evaluation of the results using a number of rating scales According to the database, it was found that it consists of several features which represent the coordinates of X, Y represent the coordinates, DMC, DC, ISI, RH, T, Wind, Rain, and FFMC.

3.2 PKC-GRU Design Stages

The effectiveness determination model consists of four stages in this section: The first stage is collecting data from the forest fires database and conducting pre-processing of the data. The database and the third stage is the use of determinants of knowledge with the algorithm (GRU) and the fourth stage is the use of a number of measures to test the validity of the choice of the predictor.

3.2.1 Data Preprocessing Stage

Before creating the predictor, that forest fires dataset required to be handled. The forest fires dataset is using the data or checking missing value is processed in the data processing stage.

3.2.2 Building Knowledge Constrains Stage

We relied on the forest fires database and based on the specific characteristics of the database and due to the possibility of training large data and in order to avoid the high computation and time complexity, it was better to build a predictor for forest fires to predict the burned area and the work was to use knowledge constraints in order to reduce the number of my data and thus reduce the calculations and we relied on Five equations showing the contribution of the meteorological factors causing forest fires, the first of which was the contribution of the daily maximum temperature, the contribution of the daily temperature range, and the effect of relative humidity, wind, and rain. The equations determine whether it is possible to build the predictor or not.

Algorithm 3.1 # Knowledge Constrains

Input: Dataset forest fires from KDD cups

Output:predicte the forest fires

1: **For** i in range(temp.value_counts().sum()

2: **IF** temp < 20

3: KC1=$1/(1+(1/5(20-((\text{temp}))^4$

4: **Else**

5: KC1=1

6: **Call build KC-GRU**

```
        End if
        IF (temp − temp¯) < 25
                KC2 = 1/(1 + (1/10(20 − (temp − temp¯))^6
          Else
              KC2=1
              Call build KC-GRU
          End if
        IF RH > 15%
            KC3=1/(1+(1/10(RH-20))^4
          Else
            KC3=1
            Call build KC-GRU
          End if
      IF rain > 0
        KC4=1/(1/(1+(rain)^3
        Else
         KC4=1
         Call build NPKC-GRU
        End if
      IF wind < 7
          KC5={1/(1/(1 + 1/12(7 − wind))^14}
        Else
          KC5=1
          Call build KC-GRU
END Knowledege Constrains
```

3.2.3 GRU

- It started from the stage of collecting data from a bank from the database and pre-processing it that if the data was selected, it fulfills the principle of checking for missing values.
- By applying the five types of knowledge constraints (KC1,KC2,KC3,KC4,KC5) in order to reduce the number of processed data in the database.
- After we apply the knowledge constraints to the database values, it is divided into two stages:
    - The training phase, which includes defining the basic parameters of GRU through the principle of try and error, and then you go to training (Train GRU and then you specify the size of the windows and then you go to training (Train LSTM)And after training (Train GRU) it determines if the error is less, it depends and gives the final result, otherwise it updates the update gate with weights.
    - Test Phase
- The advantage of (**GRU**) is that it uses less memory for processing, and the units of (**GRU**) deal with the issue of disappearing gradients, which most recurrent neural networks suffer from, Grading may become too little to have an impact on learning if it shrinks over time as it back propagates, rendering the neural network untrainable.
- The update gate and reset gate are two gates that **GRU**s utilize to address this issue. These gates can be trained to retain information from further back and determine what information is allowed through to the output. As a result, it can transfer important information down an event chain to provide more accurate predictions.

Algorithm 3.2#: KC-GRU

Input: FFD & KC // **FFD** ***Forest Fires*** *Dataset*

Output: predicte the Area

// Collect &Pre-Processing

For *each* *row in FFD*

For *each* *column in FFD*

Call Knowledege Constrains

1: **For *each*** *Iterations in FFD of burned area*

2: **Compute Update Gate**

4: **Compute Hidden state candidate**

5: **Compute Output Gate**

6: **End for**

7: **For** *each Train GRU*

8: **IF error <= Emax** **// max number of epoch and max error**

9: **Display Fin**al **Result**

ELSE

10: **Update the Two Gates with Weights**

End for

Call Evaluation

End KC-GRU

3.3 Summury

Weather prediction is a time series forecasting application in data science where predictions are made for a specific time using time series data and algorithms**,**This section will attampting to find answer of equations that shown in chapter number one.

- **How knowledge contrains can be useful in building Predicator?**
  Fire forecasts are usually generated from any data available for a given area without taking into account the cases of no fires for that forest, depending on the nature of the area. Therefore, the

determinants of knowledge had a significant impact in reducing the effort to build these systems, as the nature of the probability of the occurrence or non-occurrence of fires was determined. before building a forecaster.

- **How can we combine two technologies—Neurocomputing and knowledge contrains—to create a multi-layer model?**

  A network was characterized as containing several layers by mixing those parameters represented by one or more parameters for that layer, which increased the accuracy of the prediction model that was built, despite this, in turn, it led to an increase in the number of calculations that were performed, but it reduced the training time of the network and its access to the solution. More accurate in less time.

- **What is the beneficial result from building predictor by a combination between KC and GRU?**

  The determinants had a significant impact on increasing the accuracy of the results and the speed of training, while the network itself was distinguished by its advantages that made it very useful for predicting forest fires, as well as the possibility of working with data collected in real time**.**

Chapter Four

Implementation of (PKC-GRU)

Chapter Four: Implementation of (PKC-GRU)

4.1 Introduction

This chapter presents the key parameters and implementation of the design-related large data the predicator ***(KC-GRU)*** described in Chapter 3 using a large dataset based on forest fires, and testing the outcome of each phase of the model ***(PKC-GRU)*** divided into four stages, ***the first stage*** relates to collecting data set from forest fires, performing data pre-processing and selecting data from the base, otherwise it will perform a feature selection based on IG ,***the second stage*** it consists of building knowledge constrains and using the algorithm***(GRU)*** together includes five equations that depends on the contribution of daily maximum temperature, the contribution of Temperature range, diurnal temperature, influence of relative humidity, wind and precipitation, and the equations determine whether or not a predictor can be created. and finally, . As for ***the third stage***,building the predicator **(*KC-GRU)*** , and finally, ***the fourth stage*** is the use of rating scales through which the quality of the forest fire predictor is adopted.

4.2 Implementation of PKC-GRU

4.2.1 Description of DataSet

In this section, we present the computation of the dataset containing 516 rows and 13 columns containing the features ***X, Y, month, day, FFMC, DMC, DC, ISI, temperature, RH, wind, rain, area*** and these data are related to fires The data obtained are real, not imaginary, and captured from multiple sensors related to features associated with the forest fire database based on contributions from temperature, humidity, rain, and speed wind and obtains from the dataset ***(archive.ics.uci.edu)***

Table(4.1): description of the dataset

Data	Descripsion
X	x-axis spatial coordinate within map: 1 to 9
Y	y-axis spatial coordinate within map: 2 to 9
Month	month of the year: "jan" to "dec"
Day	day of the week: "mon" to "sun"
FFMC	FFMC index from the FWI system: 18.7 to 96.20
DMC	DMC index from the FWI system: 1.1 to 291.3
DC	DC index from the FWI system: 7.9 to 860.6
ISI	ISI index from the FWI system: 0.0 to 56.10
Temp	temperature in Celsius degrees: 2.2 to 33.30
RH	relative humidity in %: 15.0 to 100
Wind	wind speed in km/h: 0.40 to 9.40
Rain	outside rain in mm/m2 : 0.0 to 6.4
Area	the burned area of the forest (in ha): 0.00 to 1090.84

4.2.2 Result of Most Information Features (Information Gain and Entropy)

When we apply the Entropy and Information Gain to a forest fire database, the result is shown in the table below:

Table(4.2): Result of Entropy and Information Gain

Features	Entropy	IG
X	3.029	1.531
Y	2.193	1.075
FFMC	6.064	3.185
DMC	7.360	4.006

DC	7.383	4.039
ISI	6.309	3.390
Temp	7.301	3.958
RH	5.735	2.981
Wind	3.925	2.026
Rain	0.154	0.0530
Area	5.153	5.153

4.2.3 Result of Knowledge Constrains

When Applying Knowledge Constrains The basic features of the forest fires database will be subject to the constraints of knowledge that we apply with five equations, this is to determine the target (burned area) it turns out that the target is closely related to the temperature and that if the state is zero then there is no combustion, and vice versa, case 1 indicates the presence of combustion, And the table below shows is that

Table(4.3): Results of KC of on Temperature

X	Y	Temp	KC1
7	5	8.2	0
7	4	18	0
7	4	14.6	0
8	6	8.3	0
8	6	11.4	0
8	6	22.2	0

8	6	24.1	0
8	6	8	0
8	6	13.1	0
7	5	22.8	0
7	5	17.8	0
6	5	21.3	0
--	---	-----------	--------
8	6	20.1	1
4	6	28.3	1

In the case of applying the second equation to the forest fire database, the following result was obtained: -

Table (4.4): Result of KC of on Diurnal Temperature

X	Y	temp	KC2
7	5	8.2	0
7	4	18	0
7	4	14.6	0
8	6	8.3	0
8	6	11.4	1
8	6	22.2	1
8	6	24.1	1
8	6	8	1
8	6	13.1	1
7	5	22.8	1

7	5	17.8	1
6	5	21.3	1
----	---	-----------	--------
8	6	20.1	
4	6	28.3	

In the case of applying the third equation to the forest fire database, the following result was obtained: -

Table (4.5): Result of KC of on Relative Humidity

X	Y	RH	KC3
7	5	51	0
7	4	33	0
7	4	33	0
8	6	97	0
8	6	99	0
8	6	29	0
8	6	27	0
8	6	86	0
8	6	63	0
7	5	40	0
7	5	51	0
7	5	38	0
6	5	72	0

6	5	42	0
----	---	-----	-----
8	6	34	
4	6	26	

In the case of applying the fourth equation to the forest fire database, the following result was obtained:-

Table 4.6: Result of KC of on Rain

X	Y	Rain	KC4
7	5	0	0
7	4	0	0
7	4	0	0
8	6	0.2	1
8	6	0	0
8	6	0	0
8	6	0	1
8	6	0	0
8	6	0	1
7	5	0	1
7	5	0	1
7	5	0	1
6	5	0	1
6	5	0	1

----	---	--------	-------
8	6	0	
4	6	0	

In the case of applying the fifth equation to the forest fire database, the following result was obtained.

Table (4.7): Result of KC of on wind speed

X	Y	Wind	KC5
7	5	6.7	0
7	4	0.9	0
7	4	1.3	0
8	6	4	1
8	6	1.8	1
8	6	5.4	1
8	6	3.1	1
8	6	2.2	0
8	6	5.4	
7	5	4	
7	5	7.2	
7	5	4	
6	5	6.7	
6	5	2.2	
----	---	---------	
8	6	4.5	
4	6	3.1	

Table(4.8) : Apply of Knowledge Constrains to dataset

X	Y	FFMC	DMC	DC	ISI	temp	RH	wind	rain	KC1	KC2	KC3	KC4	KC5	Area
7	5	86.2	26.2	94.3	5.1	8.2	51	6.7	0	0	0	0	0	0	0
7	4	90.6	35.4	669.1	6.7	18	33	0.9	0	0	0	0	0	0	0
7	4	90.6	43.7	686.9	6.7	14.6	33	1.3	0	0	0	0	0	0	0
8	6	91.7	33.3	77.5	9	8.3	97	4	0.2	0	0	0	1	1	0
8	6	89.3	51.3	102.2	9.6	11.4	99	1.8	0	0	1	0	0	1	0
8	6	92.3	85.3	488	14.7	22.2	29	5.4	0	0	1	0	0	1	0
8	6	92.3	88.9	495.6	8.5	24.1	27	3.1	0	0	1	0	1	1	0
8	6	91.5	145.4	608.2	10.7	8	86	2.2	0	0	1	0	0	0	0
8	6	91	129.5	692.6	7	13.1	63	5.4	0	0	1	0	1		0
7	5	92.5	88	698.6	7.1	22.8	40	4	0	0	1	0	1		0
7	5	92.5	88	698.6	7.1	17.8	51	7.2	0	0	1	0	1		0
7	5	92.8	73.2	713	22.6	19.3	38	4	0	0	1	0	1		0
6	5	63.5	70.8	665.3	0.8	17	72	6.7	0	0	1	0	1		0
6	5	90.9	126.5	686.5	7	21.3	42	2.2	0	0	1	0	1		0
----	---	----------	----------	----------	----------	----------	-----	----------	--------	--------	--------	--------	--------	--------	--------

8	6	92.2	81.8	480.8	11.9	20.1	34	4.5	0	1					1
4	6	93.5	149.3	728.6	8.1	28.3	26	3.1	0	1					1

4.2.4 Result of Correlation Coefficient

The correlation coefficient is a statistical measure of how closely two or more variables are linked. the purpose of correlation of any features affected with the target, any Positive correlation suggests that these variables increase or decrease in lockstep, whereas a negative correlation means that one variable increases as the other decreases. As a result, the features affected ***temp=0.098, DMC=0.073, FFMC=0.04,*** and ***DC=0.049*** are affected positive but the features ***ISI=0.0078, RH= -0.076,rain= -0.0074*** are effected ***Negative*** and we train GRU neural network over 50 epochs.

Table :4.9: Result of Correlation Coefficient

	X	Y	FFMC	DMC	DC	ISI	Temp	RH	Wind	rain	area
X	1	0.54	-0.017	-0.045	-0.084	0.0097	-0.047	0.086	0.019	0.065	0.063
Y	0.54	1	-0.05	0.0046	-0.11	-0.028	-0.026	0.061	-0.02	0.033	0.044
FFMC	-0.017	-0.05	1	0.38	0.33	0.53	0.43	-0.31	-0.028	0.057	0.04
DMC	-0.045	0.0046	0.38	1	0.68	0.3	0.47	0.072	-0.1	0.075	0.073
DC	-0.084	-0.11	0.33	0.68	1	0.22	0.49	-0.042	-0.2	0.036	0.049

ISI	0.0097	-0.028	0.53	0.3	0.22	1	0.39	-0.14	0.11	0.068	0.0078
Temp	-0.047	-0.026	0.43	0.47	0.49	0.39	1	-0.53	-0.23	0.07	0.098
RH	0.086	0.061	-0.31	0.072	-0.042	-0.14	-0.53	1	0.07	0.1	-0.076
Wind	0.019	-0.02	-0.028	-0.1	-0.2	0.11	-0.23	0.07	1	0.061	0.012
Rain	0.065	0.033	0.057	0.075	0.036	0.068	0.07	0.1	0.061	1	-0.0074
Area	0.063	0.044	0.04	0.073	0.049	0.0078	0.098	-0.076	0.012	0.0074	1

Table(4.10) The result of Fit the model on the dataset

Iteration	Total Time	Loss
10	28s 5ms/step	0.0955
20	32sms/step	0.0832
30	44ms/step	0.0810
40	65ms/step	0.0808
50	72ms/step	0.0800

4.3 Algeria dataset is divided into two parts

Now we will compare the approved database and the Algerian database. We found that not all characteristics apply to the five approved equations ,the dataset is regression but Algerian is classification and regression, This dataset is used for educational purposes. The location of dataset https://archive.ics.uci.edu/ml/datasets/Algerian+Forest+Fires+Dataset This is part of **UCBerkeley Spring2022 Stat159 Group10's** final project The dataset includes 244 instances that regroup a data of two regions of Algeria, namely the Bejaia region located in the northeast of Algeria and the Sidi Bel-abbes region located in the northwest of Algeria,122 instances for each region.The period from June 2012 to September 2012,The dataset includes 11 attributes (12 including region attribute) and 1 output attribute (class).The 244 instances have been classified into "fire" (138 classes) and "not fire" (106 classes) classes.

Table (4.11): Description of the Algeria dataset

Data	Description
Temp	“temperature noon (temperature max) in Celsius degrees: 22 to 42”
RH	“Relative Humidity in %: 21 to 90”
Ws	“Wind speed in km/h: 6 to 29“
Rain	“0 to 16.8 FWI Components”
FFMC	“Fine Fuel Moisture Code (FFMC) index from the FWI system: 28.6 to 92.5“
DMC	“Duff Moisture Code (DMC) index from the FWI system: 1.1 to 65.9“
DC	“Drought Code (DC) index from the FWI system: 7 to 220.4“

ISI	“Initial Spread Index (ISI) index from the FWI system: 0 to 18.5“
BUI	“Buildup Index (BUI) index from the FWI system: 1.1 to 68“
FWI	“Fire Weather Index (FWI) Index: 0 to 31.1“
Classes	“Two classes, namely fire and not fire“

4.3.1 Description of Bejaia Region Dataset

This is a description of the Bejaia region database and the table below shows that,contained from 10 rows and 122 columns.

Table(4.12): Description of the Database of Bejaia Region

Temperature	RH	Ws	Rain	FFMC	DMC	DC	ISI	BUI	FWI
32	71	12	0.7	57.1	2.5	8.2	0.6	2.8	0.2
30	73	13	4	55.7	2.7	7.8	0.6	2.9	0.2
29	80	14	2	48.7	2.2	7.6	0.3	2.6	0.1
30	64	14	0	79.4	5.2	15.4	2.2	5.6	1
32	60	14	0.2	77.1	6	17.6	1.8	6.5	0.9
35	54	11	0.1	83.7	8.4	26.3	3.1	9.3	3.1

35	44	17	0.2	85.6	9.9	28.9	5.4	10.7	6
28	51	17	1.3	71.4	7.7	7.4	1.5	7.3	0.8
27	59	18	0.1	78.1	8.5	14.7	2.4	8.3	1.9
-----	-------	---	-----	------	------	------	----	-----	----
27	87	29	0.5	45.9	3.5	7.9	0.4	3.4	0.2
24	54	18	0.1	79.7	4.3	15.2	1.7	5.1	0.7

4.3.2 Result of Knowledege Constrains on Bejaia Region Dataset

It was found from the application of the equations of knowledge constraints on the Algerian database, Bejaia region,That the features of temperature and wind greatly affects the occurrence of a forest fire, and it is shown that (class) shows the state of the fire occurring or not, and the following table shows that: It is noted that the difference between (class),(class1), where it was found that (class) is from applying the knowledge constraints equations to our database, while (class1) is from applying the knowledge constraints equations to the database, After applying the first equation to the data of the Algerian forest fires database, the Bejaia region, the result was:

After applying the five equations to the data of the Algerian forest fires database, Bejaia region, the final result was:

Table(4.13) : Apply Knowledge Constrains to Bejaia Region

Temp	RH	Ws	Rain	FFC	DMC	DC	ISI	BUI	FWI	kc1	Kc2	kc3	kc4	kc5	Class	class1
29	57	18	0	65.7	3.4	7.6	1.3	3.4	0.5	1	0	0	1	0	1	0
29	61	13	1.3	64.4	4.1	7.6	1	3.9	0.4	1	0	0	1	1	1	0
26	82	22	13.1	47.1	2.5	7.1	0.3	2.7	0.1	1	0	0	1	1	1	0
25	89	13	2.5	28.6	1.3	6.9	0	1.7	0	1	0	0	1	1	1	0
27	77	16	0	64.8	3	14.2	1.2	3.9	0.5	1	0	0	1	0	1	0
31	67	14	0	82.6	5.8	22.2	3.1	7	2.5	1	0	0	1	0	1	1
33	54	13	0	88.2	9.9	30.5	6.4	10.9	7.2	1	0	0	1	0	1	1
30	73	15	0	86.6	12.1	38.3	5.6	13.5	7.1	1	0	0	1	0	1	1
25	88	13	0.2	52.9	7.9	38.8	0.4	10.5	0.3	1	0	0	1	1	1	0
28	79	12	0	73.2	9.5	46.3	1.3	12.6	0.9	1	0	0	1	0	1	0

26	32	--	28	30	27		26	31
80	47	-	80	78	84		81	65
16	14	-	17	20	21		19	14
1.8	0.7	-	3.1	0.5	1.2		0	0
47.4	77.5	-	49.4	59	50		84	84.5
2.9	7.1	-	3	4.6	6.7		13.8	12.5
7.7	8.8	-	7.4	7.8	17		61.4	54.3
0.3	1.8	-	0.4	1	0.5		4.8	4
3	6.8	-	3	4.4	6.7		17.7	15.8
0.1	0.9	-	0.1	0.4	0.2		7.1	5.6
1	1	-	1	1	1		1	1
0	0	-	0	0	0	0	0	0
0	0	-	0	0	0		0	0
1	1	-	1	1	1		1	1
1	1	-	1	1	1		0	0
1	1	-	1	1	1		1	1
0	0	-	0	0	0		1	1

4.3.3 Description Sidi-Bel Abbes region

This is a description of the Sidi-Bel Abbes region database and the table below shows that,it consisted of from 10 rows and 122 columns.

Table(4.14) : Description Sidi-Bel Abbes Region

Temperature	RH	Ws	Rain	FFMC	DMC	DC	ISI	BUI	FWI
32	71	12	0.7	57.1	2.5	8.2	0.6	2.8	0.2
30	73	13	4	55.7	2.7	7.8	0.6	2.9	0.2
29	80	14	2	48.7	2.2	7.6	0.3	2.6	0.1
30	64	14	0	79.4	5.2	15.4	2.2	5.6	1
32	60	14	0.2	77.1	6	17.6	1.8	6.5	0.9
35	54	11	0.1	83.7	8.4	26.3	3.1	9.3	3.1
35	44	17	0.2	85.6	9.9	28.9	5.4	10.7	6
28	51	17	1.3	71.4	7.7	7.4	1.5	7.3	0.8
27	59	18	0.1	78.1	8.5	14.7	2.4	8.3	1.9
30	41	15	0	89.4	13.3	22.5	8.4	13.1	10
31	42	21	0	90.6	18.2	30.5	13.4	18	16.7
27	58	17	0	88.9	21.3	37.8	8.7	21.2	12.9
30	52	15	2	72.3	11.4	7.8	1.4	10.9	0.9
27	79	16	0.7	53.4	6.4	7.3	0.5	6.1	0.3
28	90	15	0	66.8	7.2	14.7	1.2	7.1	0.6
----------	----	------	------	--------	-------	-------	------	------	----
27	87	29	0.5	45.9	3.5	7.9	0.4	3.4	0.2
24	54	18	0.1	79.7	4.3	15.2	1.7	5.1	0.7

4.3.4 Result of knowledege constrains on Sidi-Bel Abees Region

Now a region Sidi-Bel Abbes database will be applied to the knowledge constraints equations, After applying the first equation to the data of the Algerian forest fires database idi-Bel Abbes the result was

Finally, after applying the five equations to the data of the Algerian forest fires database Sidi-Bel Abbes, the result was:

Table(4.15): Apply Knowledge Constrains to Sidi-Bel Abees Region

Temp	RH	Ws	Rain	FFMC	DMC	DC	ISI	BUI	FWI	kc1	Kc2	kc3	kc4	kc5	Class1	Class2
32	71	12	0.7	57.1	2.5	8.2	0.6	2.8	0.2	1	0	0	1	0	1	0
30	73	13	4	55.7	2.7	7.8	0.6	2.9	0.2	1	0	0	1	0	1	0
29	80	14	2	48.7	2.2	7.6	0.3	2.6	0.1	1	0	0	1	0	1	0
30	64	14	0	79.4	5.2	15.4	2.2	5.6	1	1	0	0	1	0	1	0
32	60	14	0.2	77.1	6	17.6	1.8	6.5	0.9	1	0	0	1	0	1	1
35	54	11	0.1	83.7	8.4	26.3	3.1	9.3	3.1	1	0	0	1	0	1	1
35	44	17	0.2	85.6	9.9	28.9	5.4	10.7	6	1	0	0	1	0	1	0

24	27	-------	30	27	28
54	87	--	41	59	51
18	29	---	15	18	17
0.1	0.5	----	0	0.1	1.3
79.7	45.9	-------	89.4	78.1	71.4
4.3	3.5	---	13.3	8.5	7.7
15.2	7.9	--	22.5	14.7	7.4
1.7	0.4	---	8.4	2.4	1.5
5.1	3.4	--	13.1	8.3	7.3
0.7	0.2	---	10	1.9	0.8
1	1	---	1	1	1
0	0	---	0	0	0
0	0	---	0	0	0
1	1	---	1	1	1
0	0	--	0	0	0
1	1	-------	1	1	1
0	0	-------	1	1	0

It was found from the application of the equations of knowledge constraints on the Algerian database, Sidi-Abees region

The features of temperature and wind greatly affects the occurrence of a forest fire, and it is shown that (class) shows the state of the fire occurring or not, and the following table shows that, It is noted that the difference between (class),(class1), where it was found that (class) is from applying the knowledge constraints equations to the database, while (class2) is from applying the knowledge constraints equations to the database .

Table(4.16) : Evaluation(Train,Test)

	0	1	2	3	4	5	6	7	8	9	10
0	0.750	0.428571	0.870968	0.086492	0.101325	0.090909	0.192926	0.423529	0.700000	0.00000	0.0

1	0.750	0.285714	0.927742	0.118194	0.775419	0.119430	0.508039	0.211765	0.055556	0.00000	0.0
2	0.750	0.285714	0.927742	0.146795	0.796294	0.119430	0.398714	0.211765	0.100000	0.00000	0.0
3	0.875	0.571429	0.941935	0.110958	0.081623	0.160428	0.196141	0.964706	0.400000	0.03125	0.0
4	0.875	0.571429	0.910968	0.172984	0.110590	0.171123	0.295820	0.988235	0.155556	0.00000	0.0

Chapter Five:

Conclusions

Chapter five: Conclusions

5.1 Conclusion

The conclusions from this work are summarized in this chapter as follows:

A. We faced many difficulties because the database was ready and was not captured by any sensors related to meteorological factors. Therefore, we had to rely on mathematical methods to determine the burned area of forest fires, and we relied on knowledge Constraints.

B. The primary contribution of meteorological elements from forest fires occurs analysis point of view, there are certain factors that affect every single occurrence of forest fires, but forest fire causes can only be one factor in the choice. a variety of outcomes as a result of multiple factors interacting in this section. Moreover, it was proved that the (NPKC-GRU) system gives the best results according to the evaluation scales related to the area burned in forest fires. Through the application of the specified criteria, we found that (DMC, TEMP, and ISI) are the most influential for the occurrence or increase of the probability of a forest fire.

C. The GRU has two gates and does not possess any internal memory. In GRU reset gate is applied directly to the previous hidden state, GRU uses less training parameters and therefore uses less memory and executes faster GRU is used when you have less memory consumption and want faster results.

References

References

Abdi, O., Kamkar, B., Shirvani, Z., Teixeira da Silva, J. A., & Buchroithner, M. F. (2018). Spatial-statistical analysis of factors determining forest fires: a case study from Golestan, Northeast Iran. Geomatics, Natural Hazards and Risk, 9(1), 267-280, https://**doi**.org/10.1080/19475705.2016.1206629

Al-Kahlout, M. M., Ghaly, A. M. A., Mudawah, D. Z., & Abu-Naser, S. S. (2020). Neural Network Approach to Predict Forest Fires using Meteorological Data. International Journal of Academic Engineering Research (IJAER), 4(9), International Journal of Academic Engineering Research (IJAER) ISSN: 2643-9085 Vol. 4 Issue 9, September – 2020, Pages: 68-72

Basavaraju, S., Gaj, S., & Sur, A. (2019). Object memorability prediction using deep learning: Location and size bias. Journal of Visual Communication and Image Representation, 59, 117-127, https://doi.org/10.1016/j.jvcir.2019.01.008

Ben-Naim, A. (2020). Entropy and time. *Entropy*, *22*(4), 430. Department of Physical Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israe *Entropy* 2020, *22*(4), 430; https://doi.org/10.3390/e22040430

Benzekri, W., El Moussati, A., Moussaoui, O., & Berrajaa, M. (2020). Early forest fire detection system using wireless sensor network and deep learning. International Journal of Advanced Computer Science and Applications, 11(5).

Bernard, A. (2021). Solving interoperability and performance challenges over heterogeneous IoT networks: DNS-based solutions (Doctoral dissertation, Institut Polytechnique de Paris),

Bernard, D., Msigwa, C., & Yun, J. (2022). Towards IoT-based Medical Edge Devices: PPG-based Blood Pressure Estimation Application. IEEE Internet of Things Journal, 10.1109/JIOT.2022.3222477.

Burnad, D. H., Jain, K., Sutar, B., Joshi, C., & Mahajan, H. (2021). Forest fire detection and prediction using Internet of things, https://doi.org/10.46243/jst.2021.v6.i04.pp653-665

Caton-Kerr, S.E., Ali Tohidi, A., Gollner, M.J., (2019), Firebrand generation from the really degraded cylindrical wooden dowels. Front. Mech. Eng. 5, 32 12 pages, | https://doi.org/10.3389/fmech.2019.00032

Chandan, G., Jain, A., & Jain, H. (2018). Real time object detection and tracking using Deep Learning and OpenCV. In 2018 international conference on inventive research in computing applications (ICIRCA) (pp. 1305-1308). IEEE, doi: 10.1109/ICIRCA.2018.8597266.

Chandra, Y. I., Riastuti, M., Kosdiana, K., & Nugroho, E. P. (2020). Automatic Garden Umbrella Prototype with Light and Rain Sensor Based on Arduino Uno Microcontroller. *Int. J. Artif. Intell. Robot*, *2*(2), 42, DOI: 10.25139/ijair.v2i2.3152

Chen, Z., Badrinarayanan, V., Lee, C. Y., & Rabinovich, A. (2018, July). Gradnorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *International Conference on Machine Learning* (pp. 794-803). PMLR.

Chernova, E., Bychenkova, V., Kotova, N., & Pupykin, K. (2019, March). Automation of processes of temperature modes control in security system based on HACCP principles. In *IOP Conference Series: Materials Science and Engineering* (Vol. 497, No. 1, p. 012108). IOP Publishing, doi:10.1088/1757-899X/497/1/012108

Cortez, P., & Morais, A. D. J. R. (2007). A data mining approach to predict forest fires using meteorological data.

Dan Liu and Yanrong Zhang (2015), Research of Regional Forest Fire Prediction Method based on Multivariate Linear Regression, International Journal of Smart Home Vol. 9, No. 1 (2015), pp. 13-22, http://dx.doi.org/10.14257/ijsh.2015.9.1.02

Elvan, O. D., Birben, Ü., Özkan, U. Y., Yıldırım, H. T., & Türker, Y. Ö. (2021). Forest fire and law: an analysis of Turkish forest fire legislation based on Food and Agriculture Organization criteria. Fire Ecology, 17(1), 1-15, https://doi.org/10.1186/s42408-021-00102-7..

Escolano, S. (2003). On the information contents of urban layouts. In Fourth International Space Syntax Conference. London.

Firoz, A., Laxmi, G., & and Abdul, Q. (2018). Remote sensing and geographic information system based agroforestry suitability mapping and area identification in part of Jharkhand. Indian Forester, 144(4), 343-353, http://indianforester.co.in/.../126555

Gao, S., Huang, Y., Zhang, S., Han, J., Wang, G., Zhang, M., & Lin, Q. (2020). Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. Journal of Hydrology, 589, 125188

Gayathri, S., Karthi, P. A., & Sunil, S. (2022). Prediction and Detection of Forest Fires based on Deep Learning Approach. Journal of Pharmaceutical Negative Results, 429-433. ,DOI: 10.47750/pnr.2022.13.S03.071

Gnusov, M. A., Popikov, P. I., Malyukov, S. V., Sherstyukov, N. A., & Pozdnyakov, A. K. (2020), Improving the efficiency of forest fire prevention and suppression with of forest fire machine. In IOP Conference Series: Materials Science and Engineering (Vol. 919, No. 3, p. 032025). IOP Publishing, IOP Conference Series: Materials Science and Engineering, Volume 919, Mechanical engineering and automation of technological processes for Industry 4.0 Citation M A Gnusov et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 919 032025.

Gruber, N., & Jockisch, A. (2020). Are GRU cells more specific and LSTM cells more sensitive in motive classification of text?. Frontiers in artificial intelligence, 3, 40., https://doi.org/10.3389/frai.2020.00040

Guo, Y., Hou, L., Li, M., Yuan, Y., Liu, S., Xue, J., ... & Xu, M. (2022). Decision-making of Emergent Incident based on P-MADDPG. arXiv preprint arXiv:2203.12673., Yibo Guo, Lishuo Hou, Mingxin Li, Yue Yuan, Shun Liu, Jingyi Xue, Yafang Han, Mingliang Xu.Decision-making of Emergent Incident based on P-MADDPG. arxiv:submit/4215882, 2022 This paper was produced by NSFC 61602421 This paper was produced by China Postdoctoral Science Foundation 2016M600584.

Huang, T. J. (2017). Imitating the brain with neurocomputer a “new” way towards artificial general intelligence. International Journal of Automation and Computing, 14(5), 520-531., DOI: 10.1007/s11633-017-1082-y

Khan, S., & Khan, A. (2022). FFireNet: Deep Learning Based Forest Fire Classification and Detection in Smart Cities. Symmetry, 14(10), 2155, doi.org/10.3390/ sym1410215

Kukuk, S. B., & Kilimci, Z. H. (2021). Comprehensive analysis of forest fire detection using deep learning models and conventional machine learning algorithms. International Journal of Computational and Experimental Science and Engineering, 7(2), 84-94., https://doi.org/10.22399/ijcesen.950045

Naegele, M., Stoppel, K., Ridderbusch, H., & Dekorsy, T. (2020, February). Passively Q-switched Nd: YVO4 laser operating at 914 nm. In Solid State Lasers XXIX: Technology and Devices (Vol. 11259, pp. 138-145). SPIE.

Park, J. H., Lee, S., Yun, S., Kim, H., & Kim, W. T. (2019). Dependable fire detection system with multifunctional artificial intelligence framework. Sensors, 19(9), 2025, https://doi.org/10.3390/s19092025

Pham, B. T., Jaafari, A., Avand, M., Al-Ansari, N., Dinh Du, T., Yen, H. P. H., ... & Tuyen, T. T. (2020). Performance evaluation of machine learning methods for forest fire modeling and prediction. Symmetry, 12(6), 1022, https://doi.org/10.3390/sym12061022.

Rasooli, S. B., Rashidi, S., & Shabanian, N. (2021). Fire sensitivity of broadleaf tree species in plantations of Kurdistan, Iran. Journal of Forestry Research, 32(3), 1167-1176, https://doi.org/10.1007/s11676-020-01185-9

Salam, A. (2020). Internet of things for sustainable forestry. In *Internet of Things for Sustainable Community Development* (pp. 147-181). Springer, Cham, https://doi.org/10.1007/978-3-030-35291-2_5

Sazawa, K., Yoshida, H., Okusu, K., Hata, N., & Kuramitz, H. (2018). Effects of forest fire on the properties of soil and humic substances extracted from forest soil in Gunma, Japan. Environmental Science and Pollution Research, 25(30), 30325-30338, https://doi.org/10.1007/s11356-018-3011-1

Scheper, A. C., Verweij, P. A., & van Kuijk, M. (2021). Post-fire forest restoration in the humid tropics: A synthesis of available strategies and knowledge gaps for effective restoration. Science of the Total Environment, 771, 144647, https://doi.org/10.1016/j.scitotenv.2020.144647

Singh, K. R., Neethu, K. P., Madhurekaa, K., Harita, A., & Mohan, P. (2021). Parallel SVM model for forest fire prediction. Soft Computing Letters, 3, 100014, https://doi.org/10.1016/j.socl.2021.100014.

Stephenson, M., Anderson, D., Khalifa, A., Levine, J., Renz, J., Togelius, J., & Salge, C. (2020, July). A continuous information gain measure to find the most discriminatory problems for ai benchmarking. In 2020 IEEE Congress on Evolutionary Computation (CEC) (pp. 1-8). IEEE.

Tedim, F., Samora-Arvela, A., Aranha, J., Coimbra, C., Correia, F., & Pinto, D. M. (2023). Bridging knowledge gaps on the role of spatial planning in wildfire risk reduction: insights from Portugal. International Journal of Wildland Fire, https://doi.org/10.1071/WF22131

Tsallis, C. (2022). Entropy. Encyclopedia, 2(1), 264-300, *Encyclopedia* 2022, *2*(1), 264-300; https://doi.org/10.3390/encyclopedia2010018

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, ,https://doi.org/10.1155/2018/7068349

Wood, D. A. (2021). Prediction and data mining of burned areas of forest fires: Optimized data matching and mining algorithm provides valuable insight. Artificial Intelligence in Agriculture, 5, 24-42, https://doi.org/10.1016/j.aiia.2021.01.004

Wu, K., Nan, Q., & Wu, T. (2020). Philosophical analysis of the meaning and nature of entropy and negative entropy

Yang, L., and Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295-316.

Yoon, H. J., & Voulgaris, P. (2022). Multi-time Predictions of Wildfire Grid Map using Remote Sensing Local Data. arXiv preprint arXiv:2209.10102.

Zhang, G., Wang, M., & Liu, K. (2019). Forest fire susceptibility modeling using a convolutional neural network for Yunnan province of China. International Journal of Disaster Risk Science, 10(3), 386-403, https://doi.org/10.1007/s13753-019-00233-1

Zhang, N., Shen, S. L., Zhou, A., & Jin, Y. F. (2021). Application of LSTM approach for modelling stress–strain behaviour of soil. *Applied Soft Computing*, *100*, 106959,https://doi.org/10.1016/j.asoc.2020.106959

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. IEEE transactions on neural networks and learning systems, 30(11), 3212-3232.,doi: 10.1109/TNNLS.2018.2876865.

الخلاصة

تعتبر الغابات من الثروات الرئيسية على سطح الكرة الأرضية ولها دور أساسي في أنها توفر بيئة مناسبة لحياة العديد من الكائنات الحية بالإضافة إلى الحفاظ على قدرتها على الحصول على مناخ معتدل على سطح الكرة الأرضية. من ناحية أخرى ، تعتبر حرائق الغابات من أهم الكوارث التي تتعرض لها الأرض من حين لآخر بسبب عدة ظواهر منها الطبيعية والصناعية ، والتي قد تكون بسبب تدخل الإنسان أو التكنولوجيا.

تحاول هذه الرسالة بناء نموذج (توقع جديد لقيود المعرفة الوحدة المتكررة المسورة) بناءً على صياغة النماذج الرياضية للمحددات الأساسية التي قد تكون سببًا لحرائق الغابات ، حيث تم بناء نموذج مسمى (توقع جديد لقيود المعرفة الوحدة المتكررة المسورة) ، وتألفت من أربع مراحل: تضمنت المرحلة الأولى تجميع قواعد بيانات مختلفة للبحث العلمي حول حرائق الغابات ، والتي تضمنت العديد من الخصائص ، وكان من الضروري تحديد أي من هذه الخصائص أكثر أهمية في حدوثها. من هذه الحرائق في تلك الغابات ، أما المرحلة الثانية فهي تعتبر أساس النموذج المقترح والذي من خلاله تم استخدام أهم الخصائص لبناء خمسة محددات متفاوتة الأهمية وكانت على شكل صيغ رياضية تعتمد على أكثر الخصائص المهمة.المرحلة الثالثة من بناء المتنبئ الذي تم تلقيه من التدريب والاختبار ، تضمنت المرحلة الأخيرة من النموذج المقترح باستخدام خمسة مقاييس تقييم وكانت (GRU) تحدد جودة البيانات النتائج مشجعة ومقبولة ، مقارنة ببقية الشبكات ، حيث تتدرب أسرع وتعمل بشكل أفضل على بيانات تدريب أقل وبالتالي يسهل تدريبها ، حيث أنها تحتوي على بوابتين فقط: بوابة التحديث وبوابة إعادة التعيين فقط تقوم (GRU)) بتصفية أي معلومات غير ذات صلة بشكل انتقائي مع الاحتفاظ بما هو مفيد. هذه البوابات هي في الأساس متجهات تحتوي على قيم تبدأ من 0 والتي سيتم ضربها ببيانات الإدخال و / أو الحالة المخفية. تشير القيمة 0 في متجه البوابة إلى أن البيانات المقابلة في الإدخال أو الحالة المخفية ليست مهمة وبالتالي تُرجع كصفر.من ناحية أخرى ، تعني القيمة 1 في متجه البوابة أن البيانات المقابلة مهمة بينت كذلك ان قاعدة البيانات المستخدمة GRU، وعندما قارنا قاعدة البيانات المعتمدة مع قاعدة البيانات الجزائرية في منطقتيها ، بجاية وسيدي بالعبس ، أظهر أن قاعدة بيانات الجزائر لديها استجابة ضعيفة لمحددات المعرفة ، لأن وصف قاعدة البيانات المعتمدة هو الانحدار بينما قاعدة البيانات الجزائرية هي التصنيف والانحدار

جمهورية العراق

وزارة التعليم العالي و البحث العلمي

جامعة بابل

كلية العلوم للبنات

قسم علوم الحاسوب

طريقة مقترحة للتنبؤ بحرائق الغابات

رسالة مقدمة

الى مجلس كلية العلوم للبنات – جامعة بابل كجزء من متطلبات نيل درجة الماجستير في علوم الحاسوب

من قبل

نورس يحيى الخفاجي

باشراف

الاستاذ الدكتورة

سماهر حسين علي الجنابي

1444ه 2023م