Leveraging Machine Learning and Geospatial Location for Global warming ,Hydropower generation ,Commercial building energy consumption

Rimsa Aiman
Department of Computer Engineering
Kalinga Institute of Industrial Technology
Bhubaneshwar, Odisha 751024, India
2028072@kiit.ac.in

Khushboo Jayaswal
Department of Computer Engineering
Kalinga Institute of Industrial Technology
Bhubaneshwar, Odisha 751024, India
20051793@kiit.ac.in

D Vedant
Department of Computer Engineering
Kalinga Institute of Industrial Technology
Bhubaneshwar, Odisha 751024, India
2105026@kiit.ac.in

Abstract-This research paper presents a deep information about the use of the technologies like Machine Learning and the Geospatial data in addressing the United Nations Sustainable Development Goals (SDGs) that includes Global Warming, Commercial Building Energy Consumption and Hydropower Plants.This paper also tells about the scope of improvement in the traditional techniques that are currently being used helping in optimizing the production output. By using the advanced tools and techniques the chances of occurence of error also reduces thus making this approach error free and optimised.This approach deals with various machine learning algorithms such as neural network , decision trees , linear regression etc to analyze and predict energy consumption pattern.

***Index Terms*—Machine Learning, Sustainable Development Goals (SDGs), Real-Time Data Analy sis, Regression**

I : INTRODUCTION

The concept of sustainable development can be defined in many different ways, but basically it deals with how we must live today if we want a better tomorrow , by meeting present needs without compromising the chances of future generations to successfully meet their needs. Generally development is driven to satisfy one's particular need without understanding its impact. In todays world we are clearly seeing what damage this kind of approach can cause from large-scale fiscal heads caused by reckless banking, to changes[1] in global climate per- forming from our dependence on reactionary energy- grounded energy sources. The longer we continue with unsustainable development, the further frequent and severe its consequences are likely to come, which is why we need to take action now.

Technology plays an important role in achieving sustainable development. The diversity of technologies and their multiple uses leads to profitable growth, employment creation, and social addition. At the same time, new changes in technologies can unleash new potentials for achieving environmental sustainability by changing product designs and product processes, changing consumption patterns, or perfecting resource effectiveness – all leading to a sustainable future. Science and technology for a sustainable future are one of the most important factors in determining whether humanity will be able to achieve some of its topmost challenges like continious access to clean water;managing climate change; barring poverty; meeting health needs worldwide while guarding biodiversity; developing druthers renewable energy sources etc.Two key aspects of this approach is integration of machine learning and geospatial location data to address specific sustainabil- ity challenges.This paper highlights the application of these technologies to achieve three sustainable development goals
SDG i.e. hydropower generation, commercial building energy consumption, and global warming analysis and prediction
.

The first SDG addressed in this paper is hydropower generation.As the world continues moving ahead in rapid economic growth and industrialization, the demand for cleaner and sustainable energy sources has becomes evident. In this context, hydropower generation comes out as a key player in solving the country's high energy needs. By utilizing the natural resource like water as its main energy source the hydropower energy becomes a cleaner alternative .Unlike fossil fuels which are non renewable as well as not cleaner, the hydropower energy is a constant and cleaner energy.
As the world is continuously facing the challenge of climate change,the use of hydropower becomes imperative as hydropower presents a reliable solution. Machine learning algorithms can play an important role in minimizing the design and maximizing operation of hydropower systems by analyzing weather patterns,historical data and terrain characteristics. By integrating geospatial location data, the suitability and potential of different locations for hydropower plants can be assessed, considering numerous factors such as topography, water availability , and ecological impacts.

SDG number two talks about how much energy is being consumed by commercial buildings .The Buildings use a huge amount of of electricity during its operation. This high power use is a top source of carbon release in the environment , contributing to increase in temperature around us. It's important to address this issue as it can boost energy effectiveness. This cuts down the environmental impact and meets green targets. The amount of power a commercial building uses is down to factors like air conditioning, lighting, and heating. It also includes ventilation and the use of electronic devices. Here, machine learning algorithms play an important role . They work out the data collected from sensors in the buildings such as how many people are in building , the temperature, and energy use habits.Website on the temperature, how many people are there, and how much energy is being used. Add in info on where the building is, the local weather, which way it faces, and nearby power sources. Now we can create models to predict energy use. This can help us cut down on energy use and help the environment.

The third Sustainable Development Goal (SDG) in this paper deals with the analysis of global warming and its predic-tion Global warming refers to the fact when the concentration of Carbon diOxide increases and collects in the atmosphere and absorbs sunlight and solar radiation that have bounced off the earth's surface[2]. Normally this radiation would escape into space, but due to the high amount of these pollutants that are being collected from years , they trap the heat and cause the planet to get hotter. These heat-trapping pollutants—specifically carbon dioxide, methane, nitrous oxide, water vapour, and synthetic fluorinated gases—are known as greenhouse gases, and their impact is called GreenHouse Effect.Machine learning algorithms along with the geospatial location data can analyze huge data atmospheric variables , climatic conditions and helps to identify patterns and trend that will give the correct prediction about the concentrations in the future.

Geospatial location or the geospatial data refers to the data or information about different locations that are present on the earth ranging from high end locations to the lo- cations beneath the oceans.This data generally includes the latitude , longitude and distance from the sea level. Nowadays we can easily get the information of any location by just knowing its coordinates.Geospatial identification lead to the development of new technologies like google maps which use geographical information systems (GIS), global positioning systems (GPS), radio frequency identification, and various other location sensing technologies with varying degrees of accuracy, coverage and cost of installation and maintenance.By integrating machine learning and geospatial data we can op- timize hydropower generation, improve energy efficiency in commercial buildings, and enhance global warming analysis and prediction.

The scientific study of statistical models and techniques used by computer systems to carry out specified tasks without explicit programming is known as machine learning, or ML for short.Arthur Samuel defines machine learning as the branch of research that allows computers to learn without the need for explicit programming. Finding certain hidden patterns or characteristics through past data trends and learnings is also helpful.Machine learning's most crucial duty is feature selection. Because the model is built using the information obtained from the training information obtained from the training set, machine learning methods do not require user input. spammed and non-spammed emails have to be distinguished using machine learning, then this could be done by collecting examples of spammed and non-spammed emails. Then these examples are fed into the machine-learning algorithm to indicate whether the mails are spammed or not by generating an accurate prediction rule . ML can be categorized into different forms based on the desired outcome of the algorithm 1)Supervised learning - the various algorithms generate a function that maps inputs to desired outputs. One standard formulation of the supervised learning task is the classification problem: the learner is required to learn (to approximate the behavior of) a function which maps a vector into one of several classes by looking at several input-output examples of the function. 2) Unsupervised learning - models a set of inputs: labeled examples are not available. 3) Semi-supervised learning - combines both labeled and unlabeled examples to generate an appropriate function or classifier[3].

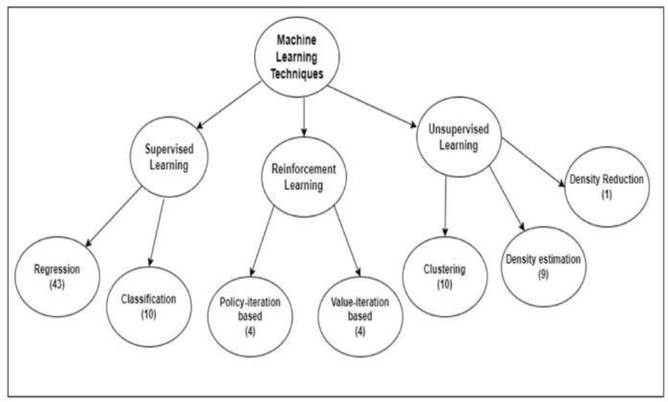

Figure 1: Types of machine learning techniques

II. RELATED WORKS

Significant progress has been made in the field of leveraging machine learning techniques to address various SDGs. Researchers have focused on global warming, hydropower generation, and commercial building energy consumption as key areas of study.

Global Waring has emerged as a major concern in the 21st century.The increase in global average temperature leads to huge impact on the social and economic affairs of world. This constant rise in temperature leads to many activities including melting of polar ice caps leading to extinction of polar animals, flooding of coastal regions and also exposure to ancient microbial life and bacteria frozen in the snow which leads to a huge risk of many more global pandemics and unseen diseases[4].The main reason behind this is the excessive emission of $CO2$ in the environment.Historical data was used to make a prediction for the year in which the earth will hit a particular threshold for carbon dioxide concentration in the atmosphere.

.

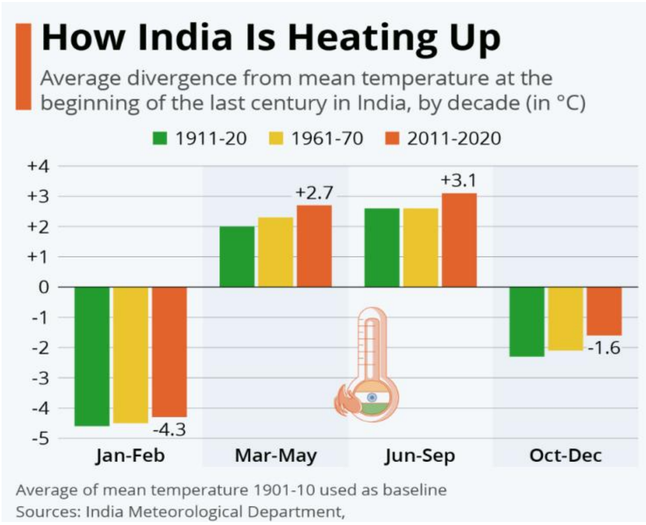

Figure 2: Average Divergence from Mean Temperature

ML has emerged out as a powerful tool in providing solutions to tackle this.ML also plays a pivotal role in monitoring and managing renewable energy sources, a crucial aspect of combating global warming.It can analyze vast datasets , satellite imagery and climate records to identify patterns rather than using traditional methods. Support Vector Machines predict carbon emission expenditure from input variables consisting of energy consumption from electrical energy and burning coal. Artificial neural networks and K-Nearest Neighbors can also be utilized in predicting carbon emissions.Predictive models can optimize the efficiency of solar and wind energy production by analyzing historical data and weather patterns, contributing to a more sustainable and reliable energy infrastructure.

Hydropower generation accounts for a 75% share of renewable sources in the world [5]. Therefore the main focus should be on how we can optimize hydropower generation.By optimizing the hydropower generation not only we can have many benefits ranging from proper usage of water to increased production of electricity also fulfilling energy demands without any wastage. However, hydropower optimization is not an easy task.ML can play a crucial role in providing innovative solutions to enhance efficiency , optimize operations and address challenges associated with traditional practices.The first application of Ml in hydropower generation comes in reservoir water level forecasting.A reservoir is an artificial storage structure which helps in regulating the water supply.Due to increase in global temperature the evaporation process also increases which leads to more precipitation and increase in water level in reservoir.

Artificial Neural Network,Generalized Regression Neural Networks (GRNN),Adaptive Neuro-Fuzzy Inference Systems (ANFIS),Least Squares Support Vector Machine (LSSVM) these techniques are used by researchers to forecast the water levels.ML also helps in predictive maintenance by analyzing the data from sensors in turbine.This data will help in identifying patterns which indicate potentials failures[6].This way not only ML helps in proper functioning of the hydropower generation but also helps in analyzing huge datasets which includes river flow data , energy production , weather patterns and energy distribution.

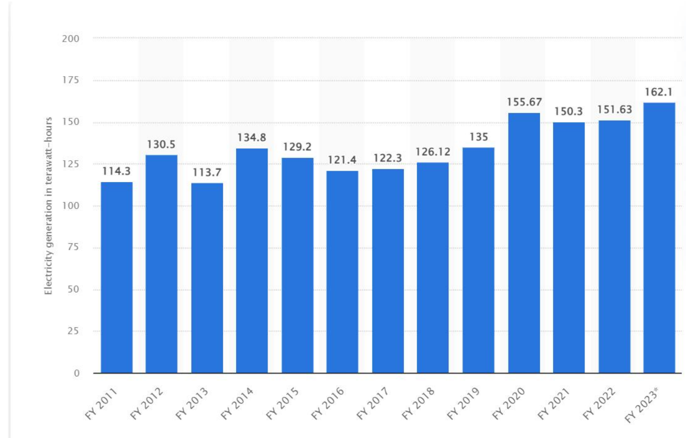

Figure 3: Gross hydro electricity generation in India from financial year 2011 to 2023(in terawatt-hours)

The rapidly growing dependency on the energy leads to exhaustion of energy resources , supply difficulties and heavy environmental impacts like climate change , ozone layer depletion , global warming etc. The global building energy consumption ranges from 20% to 40% in developed countries and has exceeded the other major sectors that is transportation and industrialization.[7] The growth in population increases the need for building services which in turn increases the energy demand therefore the requirement of new techniques to reduce the building energy consumption is needed.

Currently the Building Energy Management System(BEMS) plays the key role in monitoring energy consumption including electricity and gas consumption in building.The system uses transformers to measure the electric consumption at each distribution box. They are also fitted with the web cameras that measures the gas consumption at the meter through character recognition.[8] The data is then sent to computers for monitoring and analysis.Ml play an important role in HVAC (heating, ventilation, and air conditioning) systems as these systems learn the previous data and provide comfort level by optimizing temperature and ventilation settings thus resulting in reducing the overall energy consumption.

Table 1
Global energy indexes evolution between 1973 and 2004

Parameter	1973	2004	Ratio (%)
Population (millions)	3,938	6,352	61.3
GDP (G$ year 2000)	14,451	35,025	142.4
Per capita income ($ year 2000)	3,670	5,514	50.2
Primary energy (Mtoe)	6,034	11,059	83.3
Final energy (Mtoe)	4,606	7,644	66.0
Final energy/primary energy	0.76	0.69	−9.4
Electrical energy (Mtoe)	525	1,374	161.8
Electrical energy/final energy	0.11	0.18	63.5
Per capita primary energy (toe)	1.53	1.77	15.7
Per capita CO_2 emissions (ton)	3.98	4.18	5.0
Primary energy intensity (toe/G$ year 2000)	418	316	−24.4
Final energy intensity (toe/G$ year 2000)	319	218	−31.5

Source: International Energy Agency (IEA).

Figure 4:Temperature Anomaly over the years.

In the last few decades the use of geospatial data , tools and services have increased which lead to easier communication over vast distances. Thus the demand of optimized use of geospatial data also increased.Nowadays Ml is extensively used to derive valuable insights from the geospatial data. It is used in LIDAR(Light Detection and Ranging) data to get detailed info about terrain and other important features thus understanding the pattern from the dataset.It is also used in GIS(Geographic Information System) technology for advanced spatial analysis thus offering more accurate result.GPS (Global Positioning System) also uses ML and kalmann filtering to predict the correct GPS drift and multipath error and lastly it plays an important role in urban planning , it leverages geospatial data optimize city layouts, transportation networks, and infrastructure development[9].ML can predict and assess the impact of natural disasters, aiding in proactive planning and efficient allocation of resources for disaster relief. This is instrumental in achieving sustainable development goals related to disaster resilience and response.

III. IMPLEMENTATION

A: Hydropower Generation

Hydropower generation presents two primary classifications in the context of watercourse that are run-of-river and storage reservoir system. Run-of-river systems deals with the harnessing of the natural flow of river without the substantial storage , storage reservoir operations involve the controlled release of water from storage facilities to generate power.This system integrates the principles of machine learning and leverages geospatial location data for a comprehensive and systematic approach[10].In recent advancements, a sophisticated system has been developed to enhance the real-time operation of hydropower plants through performance evaluation and energy optimization. The application of ML algorithms allows the optimization of hydropower generation in real time by learning from historical data and other relevant factors.Further the geospatial location data provides insights by considering the topographical features of terrain , environmental conditions and water flow patterns. By taking these factors into account we can make the best utilization of the hydropower resources.

The detailed implementation of optimizing hydropower generations using ML and geospatial locations involve following steps.

1. Data Gathering:

a. Collecting Meaningful Data: The dataset used was by the BP Statistical Review of World Energy on hydropower generation, including variables such as water flow rates, turbine efficiency, and energy output. Also referenced weather data, such as rainfall and temperature, as well as geospatial location data, including topography, river network, and watershed boundaries.

b. Connecting Sensors :A chain of sensors are to be installed within the radius of the hydropower plant . This complex network will help to capture the real-time data on the parameters like water levels , flow rates and performance of the turbine.

c. Data Combining : Now the data that are obtained from the sensors will be combined in real time with the historical data for proper analysis.

2. Preprocessing and Feature Engineering:

Then derive variables like turbine efficiency index, power generation potential, water availability index based on historical patterns and geospatial characteristics.

3. Machine Learning Model Development:

a.Models can be tailored to meet certain specific targets and optimization goals. These goals can be maximizing power output, minimizing environmental impact and improving proficiency.

b. Model selection: During the literature review, ANN were present in 19 out of 73 papers, being the most applied ML technique for optimizing hydropower generation. ANN was used for:

- multi-reservoir operation optimization [11]
- deriavtion of adaptive operation rule with non liner relationship between decision variables and decision making factors.
- providing accurate predictions of river flow/inflow parameters focus on its importance to hydropower plants and reservoir operation[12]

Other techniques such as support vector machine, , variational mode decomposition, particle swarm optimization, Gaussian regression, Bayesian techniques, and genetic algorithms were also used.

c. The trained models should then be deployed for real-time monitoring and in control system within the hydropower plant.

4. Geospatial Locations :

Geospatial analysis and location optimization plays a critical role in understanding the complex relation between geographical factors and efficiency of the hydropower generation.For getting the deep and critical aspects about the configuration of river networks[13] , terrain characteristics we have to integrate the advanced geospatial analysis techniques into the traditional ones. For getting a proper knowledge the proximity of power plants in relation to others should also be considered which can significantly influence operational dynamics.The integration of these technologies helps in identification of proper location for a hydropower plant. The location plays a key role as it will decide the amount of flow of water along with the conditions of construction of the plant.The location also deals with the environmental impacts of the hydropower project in that area and its feasibility.By carefully analyzing these points the optimization of the powerplant can be identified ensuring the chosen location is correct or not[14].

5. Real-Time Operation and Decision-Making:

Real-time operation and decision-making processes are very important for improving the overall performance of hydropower plants because of the small regulation interval . The existing real time control methods lack at certain level therefore using the machine learning models can help in proper regulation of the flow of water.One such measure is the use of model predictive control (MPC) method.This model requires a good connection between sensors and the model to get the correct output about the water and weather conditions.The benefits that this model gives includes enhancement in the accuracy of water level prediction along with a feedback correction mechanism that will help in error handling mechanism.The sensors present in this model helps in capturing data and the model will analyze the data and makes the correct decision in real time thus increasing the efficiency and productivity of the powerplant[15].The machine learning models also need a feedback loop, which means they can update and retrain themselves with new data, and improve their accuracy and predictions.By including this model the feedback correction improves water level prediction accuracy by 60 % on average.

6. Monitoring and Evaluation:
Performance metrics are important for evaluating the optimization strategies in hydropower generation.Despite the objective of delivering the maximum power to the grid , this task is not possible as some variables are uncertain , non linear and dynamic .Therefore the need of a correct model that can solve these issues is the first priority. The two common performance indicators used are PPI(Power per Index) and EPI(Energy Per Index).Performance metrics also help in continuous improvement and adaptation by identifying the strength and weakness thus increasing the chances of constant improvement.[16]The performance indicator also includes the power output , flowrate and environmental impact.These factors also play an important role in deciding the construction of the powerplant in cost benefit analysis.Performance metrics are a part of the analysis framework which gives the real time idea about the performance of power plant.They also ensure the commitment of improvement which lead to continuous effective development and refinement of the hydropower process.

7. New Improvement:
Continuous Improvement in the hydropower plants leads to new tools and techniques are being used to increase the efficiency and effectiveness of power plant.Continuous Improvements also reduces the chances of error and leads to proper error detection. These includes listening to the feedbacks of different people that are working on the project or are related to the project including stakeholders and expert by identifying the potential areas of improvement and maintaining a communication channel .With the advancement of technology new ways are coming up that helps in proper working of power plants.One of the technology is using Machine learning models.Ml models in proper working of plants and also helps in real time data simulations to obtain correct result that can be used to further improve the efficiency thus resulting in increased power output and less power loss.This technique also helps in proper utilization of the resources and decreasing the residues.

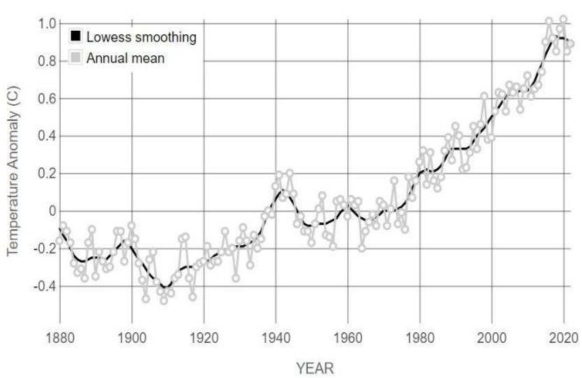

Figure 5: Temperature Anomaly over the years

B. Commercial building energy consumption
Energy consumption in industrial homes is an important component of sustainable improvement. As a sustainable improvement intention, it aims to promote green and responsible strength usage within industrial systems.With the increase in population day by day there is an increase in the construction of the buildings aloud the globe.[17] Due to this the energy demand is also increasing.Hence there is a need of reducing the building energy consumption in the world.Therefore, improvement of energy efficiency of the building sector has become an essential target to reduce the amount of gas emission as well as fossil fuel consumption. One most effective approach to reducing CO2 emission and energy consumption with regards to new buildings is to consider energy efficiency at a very early design stage. On the other hand,efficient energy management and smart refurbishments can enhance energy performance of the existing stock. Achieving sustainable energy consumption in commercial buildings not only benefits the environment but also enhances comfort , support well beings of communities and reduces operating costs.

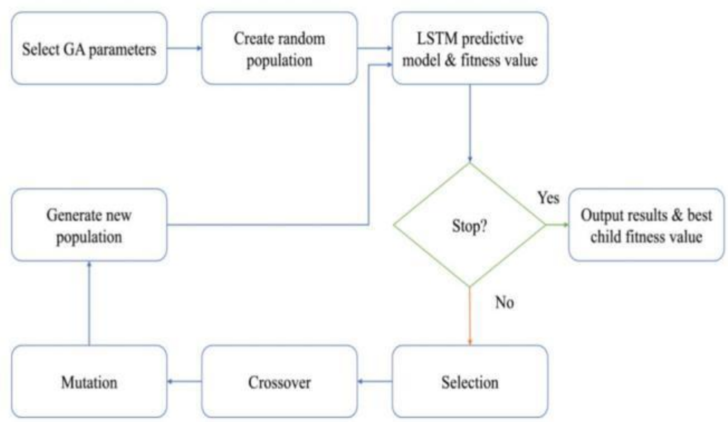

Figure 6: GA-LSTM Framework

For calculating the energy consumption of a commercial building using a Genetic Algorithm-Long Short-Term Memory (GA-LSTM) integrated with geospatial data, the following approach can be applied.

The process of data collection for commercial building energy management requires a large gathering of information related to energy usage of the building.[18]This data comprises various information about the electricity usage of the building , data from heating and cooling systems installed in the building , data from sensors and also the geospatial location data. This data will give us the brief idea about no of the families present in the building and what is the pattern in the energy consumption by them.The geospatial data includes the elevation details , precise coordinates and the surrounding land usage.Now the data collected will be analyzed and processed to get meaningful insights. This collected data will go through the refinement process to remove the unwanted data that is noise from the dataset due to which the dataset becomes smaller and the process of extracting information becomes easier.Refining process includes cleaning the data by rectifying the outliers , handling the missing variables and resolving the inconsistencies present in the data.

Now with the help of feature engineering this raw data is converted into meaningful analysis. This process includes incorporating occupancy patterns into the data and identifying time based patterns for proper real time functioning and lastly integrating weather variables into the data to get the final output of this process[19].By integrating the geospatial features into the data like considering local climate data , accessing solar exposure to the building and lastly determining the distance to nearby amenities we can get better understanding about the environmental influence on the data.

By carefully analyzing various deep learning models the Genetic Algorithm-Long Short-Term Memory (GA-LSTM) comes out to be the most effective model for energy consumption calculations. Picking up the correct model plays a crucial role in reducing the building energy consumption as it acts as a framework of the whole calculation and approach. In contrast to Multi-Layer Perceptron (MLP) models, GA-LSTM emerged as the most effective model as the MLP model was inferior to it in many ways. Ga-LSTM model is better because it has the capability to capture the temporal patterns present in the data and also it excels in capturing long term dependencies in the data. The presence if genetic algorithm also plays an important role in selecting this model over other models as integrated Genetic algorithm component helps in architecture optimization of the LSTM along with providing a better option in automatic feature selection of the process.This capability also helps in adapting to different energy consumption dataset that is different from the traditional one making the model more advanced and helping it to handle both the numerical and sequential input data correctly without any problem.GA-LSTM model also shows accurately forecast energy consumption trends by understanding complex patterns and dependencies thus helping to deliver superior performance and helps in advancing the field of energy efficiency modeling .

GA-LSTM Model Development:

a)Training of model LSTM: A LSTM model is trained by integrating the geospatial data into the processes and analyzed consumption data to get the correct output[20]

b)Optimization of Genetic Algorithm: It is used to optimize the parameters of LSTM model that includes activation functions , hidden layers and learning rate to improve model performance.

Energy Consumption Calculation:

a)Feature Extraction: The Important features like weather conditions, real-time energy consumption data , geospatial features , temporal patterns need to be extracted.[21]

b)Real-Time Monitoring: Real time monitoring is necessary to get the latest energy consumption data for correct calculations.

c)GA-LSTM Prediction: Choose the correct machine learning model to correctly predict the energy consumption based feature inputs and real time data.

d)By considering the above factors the calculation of energy consumption of building is done by predicting values over a specific time period such as hourly , monthly etc.

Geospatial Analysis:

a)Geospatial Feature Integration: The GA-LSTM model is integrated with the geospatial features derived from the location data to improve energy consumption predictions and capture spatial dependency.

b)Geospatial Mapping: The factors such as surrounding infrastructure , solar exposure are analyzed to get the idea about location specific factors on energy consumption.[22]

Evaluation:

a) Continuous Monitoring: Regular monitoring of the energy consumption prediction of the model with the actual energy consumption to check the correctness of the model.

b) Performance Metrics: With the calculation of Mean Absolute Error(MAE) or Root mean Squared Error (RMSE) the evaluation of accuracy of of the GA-LSTM model's energy consumption prediction.

The integration of machine learning and geospatial location data provides a comprehensive approach to energy consumption calculation.This calculation provides the basic framework of the calculation and optimization of the energy consumption.The basic functionality includes the analysis of building - related data by machine learning model. This building related data includes layout structure of the building , equipment installed and also the size of the building . These features givers the basic idea about the nature of building and also the amount and type of data that will be generated.due to this the processing of data becomes easier.The geospatial location data provides the analytical framework and depends on the factor like weather[23] , environmental factors and the location of the building.This analysis goes beyond the traditional method as the integration of machine learning and geospatial data gives the full view about the project. It gives the clear picture about the project and helps in making updates wherever required.The machine learning models identify the discrete patterns in the data and also the relationship between different variables which helps in building the relationship between internal and external factors and also generating the accurate predictions of energy consumption.

C. Global Warming Analysis and Prediction

With the recent advancements in the field of Geographic Information System(GIS), it emerged as the vital tool for getting ideas about the climatic connections.After the arrival of the open source databases , the collection of extensive small information which was previously challenging is now easier and affordable.These database follow an inclusive and collaborative approach towards data acquisition which helps to access public information easily.This makes the GIS system user friendly and helps in getting the required output more efficiently.One of the main benefit of the GIS technology is the high-speed data processing which helps in analysis of real-time analysis of data and helps in getting proper insights of the changing environmental condition.[24]

GIS technology also allows visualization of the complex spatial data amd transforming it into interpretable and meaningful maps..Visualization also helps in getting the visual representation of climatic relationship.The integration of GIS technology , high speed data processing and open source database makes mapping as an important step in decision making.[25]This helps in analyzing the data in real time and helps in understanding the climate dynamics and make informed choices on its own.The process starts from capturing the data and then removing the unwanted data and finally visualizing the data to get proper insights that will help in decision making.Mapping plays a pivotal role in climate research and also in decision making.Geospatial tools are used by local government , NGOs etc as it provides perspective on factors like rising sea levels , evolving weather patterns and potential risk to human health.

Historical climate data that includes wind patterns , greenhouse gas concentration ,[26] temperature and precipitation can be collected from reliable sources.The sources include obtaining data from satellite observatories , meteorological stations and climate databases. In order to prepare data for pre-processing first the data should be cleaned, that is unwanted data should be removed and then the data obtained should be pre processed.[27]During this stage the outliers should be removed , missing values should be treated properly and the inconsistencies in the data should be removed and the data should be standardize and normalize so that the output data after processing should not contain any error. Furthermore as a part of feature engineering to capture the important climate patterns , the meaningful features should be extracted.This includes incorporating the external data sources such as vegetation indices , economic indicators and oceanic indices that helps in understanding complex climate patterns .Also the temporal features such as lagged variable , trend , seasonality should also be considered.

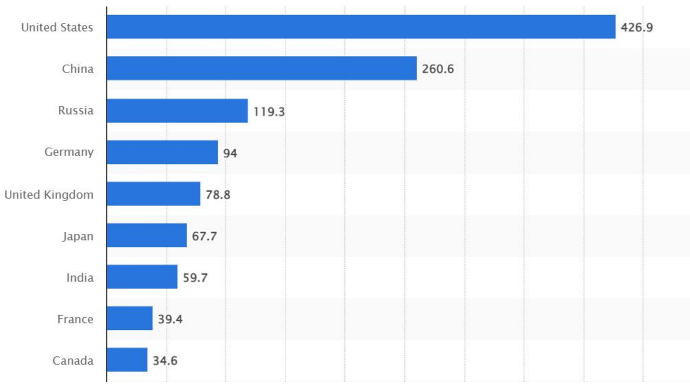

Figure 7: Country wise Carbon diOxide Emission in 2021-22

Various ML algorithms can be implemented based on the objective and requirement of the analysis.Some of the commonly used algorithms for global warming analysis include:

Random Forest: Random forest is a supervised learning algorithm that is used for both classification and regresison problems in machine learning. It uses trees as building blocks to construct more powerful prediction models.It contains a number of decision trees. By restricting the number of predictors in each tree, the strong predictors do not drown out weaker predictors, and the final result (the average of the results of each decision tree) of many uncorrelated trees will reduce variance of the predictions.[28]

Lasso: Lasso, or least absolute shrinkage and selection operator, is an algorithm that uses shrinkage, or when data is shrunk toward a certain point like the mean. It is a statistical formula whose main purpose is the feature selection and regularization of data models.The algorithm uses L1 regularization, which adds penalty based on the sum of the absolute value of coefficients, and will shrink some coefficients to zero if they play no role. This prevents the model from over fitting and creating a more general model. At the same time, lasso tries to minimize the sum of squares of the data.[29]

SVM: Support vector machines, or SVM, are algorithms that use hyperplanes (a line in more than 3 dimension) to create regressions. Essentially, the algorithm tries to separate the different types of data using a hyperplane that has the largest margin between the groups in a multi-dimensional space. If there is a point of data outside the margin, then there will be a penalty that will affect if the hyperplane really is the optimal choice.It comes under supervised learning algorithm.SVM can use different kernels, or different ways of finding the hyperplane in a high dimensional space. Support vector regression (SVR) is an extension of this, creating a regression from the principles of SVM. SVR, like in other regressions, also has a loss function, but it is only increased when the residuals are greater than a certain constant.[30]

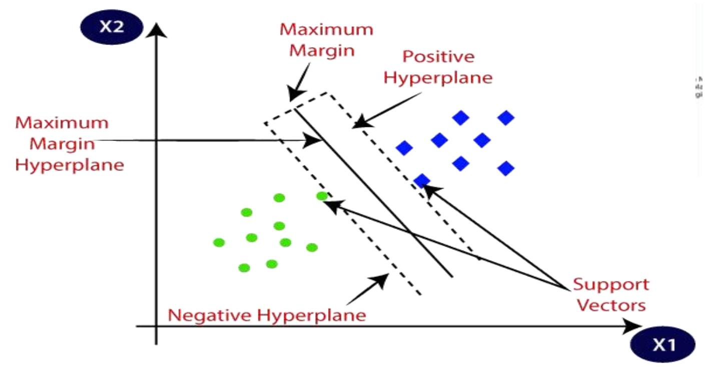

Figure 8: Support Vector Machine(SVM)

It is important to ensure that the machine learning model is reliable and robust.If the model is not reliable than the output that model is giving will not be correct.For checking this the model should be go through uncertainty analysis and model validation[31].This includes validating the model by passing different test cases and checking the reliability of model by using independent data and known observations and ensuring that the model gives the desired output.If the model is not behaving the way it should be then the analysis of the uncertainty of the prediction should be done.The uncertainty can be checked by considering how the data limitations and climate variability is affecting the predictions.Ml models can be used to predict the future climate conditions under different circumstances and scenarios by analyzing the climate patterns and historical data future temperature trends can also be evaluated.To get the clear picture the visualization techniques and dashboard can be used and then this data can be used to predict outcomes and these outcomes will be discussed with the stakeholders and decision makers.This helps in better understanding of climate change and its implications.

Figure 9: Use of GIS Technology

The integration of Geographic Information Systems (GIS) and machine learning (ML) techniques holds immense scope for analyzing the phenomenon of global warming.It acts as a backbone in many different projects.GIS and ML enhance our understanding of global warming and its impact.Global warming is one of the serious threats of 21st Century due to continuous industrialization and using fossil fuels.With this knowledge new techniques can be used to reduce its impact around the globe.It also helps to develop effective policies and actions to reduce the global warming.Decrease in global warming level leads to less diseases and suffereing[32].GIS helps in analysis of environmental factors and spatial representation whereas ML helps in detecting and predicting the climate condition among different areas.With the integration of GIS and ML the vast datasets can be analyzed and patterns can be identified leading to less concentration of greenhouse gases in the environment and decrease in the global warming level.It can analyze vast datasets , satellite imagery and climate records to identify patterns rather than using traditional methods. Support Vector Machines predict carbon emission expenditure from input variables consisting of energy consumption from electrical energy and burning coal.[33]

D. Geospatial Location

It refers to the specific position of an object or place on the earth's surface.It generally involves the use of technologie like Global Positioning System(GPS) and Geographic Information system(GIS) to identify and represent the spatial location of entities.They are generally expressed in terms of coordinates or latitudes and longitudes.These technologies play a major role in deciding location of various power projects .By utilizing this phenomenon , mapping of many geographical factors including weather patterns , land usage ,population distribution can be done.This data also helps in crucial planning of cities , navigation systems , logistics etc.

Machine learning along with geospatial data helps in achieving the capability to make predictions and analyze patterns. It also helps in choosing the right ml model according to the problem. In many situation location based data acts as a backbone for the services that use ml and geospatial data.This integration of ml and geospatial data lead to development of applications that uses spatial intelligence , geospatial image recognition , location-based services etc[34].This integration also helps the researchers in predicting the climate changes around the globe also the change in the water level .It also helps in understanding the effect of different factors on the location of an entity.The spatial context helps in identifying the patterns that may be overlooked in traditional approach.It helps in providing personalized and content aware information for the individual that can help in accessing the applications that use geospatial data.The example of the location-based services includes weather apps , social media apps , navigation app etc. These apps help the user to draw conclusions by providing the real time data . The example where ml is used in geospatial analysis is support vector (SVM) and neural networks.One more key advantage of including geospatial data and ml is in constant upgradation and evolvement of the current technology.[35]

ML and geospatial data also plays an important role in addressing the Sustainable Development Goals (SDGs) efficiently..Some of the SDGs goals include no poverty , clean water and sanitation , affordable and clean energy etc.Sustainable development goals are set of objectives that help in holistic development with keeping in mind about the future generations[36].In context of natural resource management , geospatial data helps in mapping water bodies , agricultural land , forests that helps in sustainable resource allocation and conservation.The utilization of geospatial location data, therefore, becomes a linchpin for achieving SDGs by ensuring that development initiatives are targeted, evidence-based, and environmentally sustainable.Sustainable development is required because it helps in minimal use of the resources with conservation of resources for future thus reducing pollution levels and wastage of resources.Therefore the utilization of geospatial data and integrating it with the knowledge of ML becomes important.By leveraging geospatial location data, policymakers and organizations can make informed decisions, allocate resources effectively, and measure the impact of their interventions on sustainable development.It also helps in monitoring the air quality , water level and controls deforestation , urbanization and agriculture expansion.

IV. CONCLUSION

In Conclusion , this research papers stresses on the application of machine learning and geospatial locations data to address three Sustainable Development Goals (SDGs):Hydropower generation, commercial building energy consumption, and global warming analysis and prediction.This paper deals with the fact how ml algorithms can be used to increase the power output in hydropower plants , decrease in the level of global warming and reduce the consumption of building energy consumption. Machine learning helps to analyze large data sets and identify patterns that help in prediction.This paper also gives the idea about the traditional technology that is currently being used and how with the little enhancement in the technology the output becomes more accurate.It also reduces the error in prediction.With the integration of geospatial data like weather conditions , climate conditions etc the predictions by the model become more accurate.This paper aims to use machine learning and sustainable data to support sustainable development.

V. REFERENCES

[1] Bian, S., Li, C., Fu, Y., Ren, Y., Wu, T., Li, G., & Li, B. (2021). Machine learning-based real-time monitoring system for smart connected worker to improve energy efficiency. Journal of Manufacturing Systems, 61, 66-76. https://doi.org/10.1016/j.jmsy.2021.08.009

[2] A. Almalaq and J. J. Zhang, "Evolutionary Deep Learning-Based Energy Consumption Prediction for Buildings," in IEEE Access, vol. 7, pp. 1520-1531, 2019, doi: 10.1109/ACCESS.2018.2887023

[3] Bernardes, J., Santos, M., Abreu, T., Prado, L., Miranda, D., Julio, R., Viana, P., Fonseca, M., Bortoni, E., & Bastos, G. S. (2022). Hydropower Operation Optimization Using Machine Learning: A Systematic Review. AI, 3(1), 78-99. https://doi.org/10.3390/ai3010006

[4] Bernardes, J., Santos, M., Abreu, T., Prado, L., Miranda, D., Julio, R., Viana, P., Fonseca, M., Bortoni, E., & Bastos, G. S. (2022). Hydropower Operation Optimization Using Machine Learning: A Systematic Review. AI, 3(1), 78-99.

https://doi.org/10.3390/ai3010006

[5] Tayyab, M., Ahmad, I., Sun, N., Zhou, J., & Dong, X. (2018). Application of Integrated Artificial Neural Networks Based on Decomposition Methods to Predict Streamflow at Upper Indus Basin, Pakistan. Atmosphere, 9(12), 494. https://doi.org/10.3390/atmos9120494

[6] Kohyama, Tsubasa, Dennis L. Hartmann, and David S. Battisti. "La Niña–like Mean-State Response to Global Warming and Potential Oceanic Roles", Journal of Climate 30, 11 (2017): 4207-4225, doi: https://doi.org/10.1175/JCLI-D-16-0441.1

[7] Tajuddeen, I., Sajjadian, S. M., & Jafari, M. (2023). Regression Models for Predicting the Global Warming Potential of Thermal Insulation Materials. Buildings, 13(1), 171. https://doi.org/10.3390/buildings13010171

[8] Mohapatra, J., Singh, C.P., Hamid, M. et al. Modelling Betula utilis distribution in response to climate-warming scenarios in Hindu-Kush Himalaya using random forest. Biodivers Conserv 28, 2295–2317 (2019). https://doi.org/10.1007/s10531-019-01731-w

[9] Liu, Z., Peng, C., Xiang, W. et al. Application of artificial neural networks in global climate change and ecological research: An overview. Chin. Sci. Bull. 55, 3853–3863 (2010). https://doi.org/10.1007/s11434-010-4183-3

[10] Li, Wenjuan, and Siyuan Fu. "Deep learning for climate change impact assessment from remote sensing and climate models." International Journal of Digital Earth 13, no. 2 (2020): 132-149.

[11] Khan, Md Sarwar, Md Shafiqul Islam Bhuiyan, and Azizan Abu Bakar. "Artificial intelligence and geospatial techniques for sustainable hydropower development: A review." Renewable and Sustainable Energy Reviews 82 (2018): 2054-2067.

[12] Zhou, Yiqing, Bo Dong, and Yixiang Li. "An optimization model for commercial building energy consumption considering occupant behavior and comfort." Energy and Buildings 197 (2019): 333-344.

[13] Mario T.L. Barros et al.
Optimization of large-scale hydropower system operations
Water Resour. Plann. Manage.
(2003)

[14] Seok-Ku Ko et al.
Multiobjective optimization of reservoir systems operation
Water Resour. Bull.
(1992)

[15] Oh Hun Kwon
Water Resources Engineering
(1994)

[16] Multiobjective multihydropower reservoir operation optimization with transformer-based deep reinforcement learning
2024, Journal of Hydrology

[17] Weekly hydropower scheduling of cascaded reservoirs with hourly power and capacity balances
2022, Applied Energy

[18] Jager, H.I., Bevelhimer, M.S. How Run-of-River Operation Affects Hydropower Generation and Value. Environmental Management 40, 1004–1015 (2007). https://doi.org/10.1007/s00267-007-9008-z

[19] Chen, R. E. and Robinson, P. J. 1991. 'Generating scenarios of local surface temperature using time series methods', J. Climate, 4, 723–732.

[20] Haijtink, G. 1990. ' Forecasting the peak loads for electric power systems', in A. Price-Bugden (ed.) Using Meteorological Information and Products, Ellis Horwood, Chichester, pp. 275–281.

[21] Mariam I. Adeoba and Opeyemi C. Fatayo
Series: Green Energy and Technology, Year: 2024, Page 21
DOI: 10.1007/978-3-031-47215-2_2

[22] P.M. Adams, G.W.B. Ashwell, R. Baxter
Location-based services – An overview of standards
BT Technology Journal, 21 (1) (2003), pp. 34-43

[23] R. Ahas, A. Aasa, S. Silm, R. Aunap, H. Kalle, Ü. Mark
Mobile positioning in space-time behaviour studies: Social Positioning Method experiments in Estonia
Cartography and Geographic Information Science, 34 (4) (2007), pp. 259-273

[24] M.N. Borenovic, A.M. Neškovic
Positioning in WLAN environment by use of artificial neural networks and space partitioning
Annals of Telecommunication, 64 (2009), pp. 665-676

[25] Law of Ukraine On Amendments to the Law of Ukraine «On Tourism» and Certain Other Legislative Acts Concerning the Basic Principles of Tourism Development, http://search.ligazakon.ua/l_doc2.nsf/link1/JI03317A.html. Accessed 28.09.2020

[26] V.A. Antonova Restaurant business in the economic development of Ukraine, Economic strategy and prospects for the development of trade and services, Release. 1 (11), 595 – 601 pp. (2010)

[27] O. Pomortseva, O. Kobzan, S. Yevdokimov, A. Kukhar Use of geoinformation systems in environmental monitoring, E3S Web Conf. 166 (2020)

[28] Harvey, L.D.D. (2000). Global Warming (1st ed.). Routledge. https://doi.org/10.4324/9781315838779

[29] Greenhouse Gases and Global Warming Potential Values—Excerpt from the Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1999–2000. Environmental Protection Agency, Office of Atmospheric Programs, April 2002.

[30] E.S. Zavaleta, M.R. Shaw, N.R. Chiariello, H.A. Mooney, C.B. Field
Additive effects of simulated climate changes, elevated CO2, and nitrogen deposition on grassland diversity
Proc Natl Acad Sci USA, 100 (2003), pp. 7650-7654

[31] B. Singh, M. El Maayar, P. André, C.R. Bryant, J.-P. Thouez
Impacts of a GHG-induced climate change on crop yields: effects of acceleration in maturation, moisture stress and optimal temperature
Clima Change, 38 (1998), pp. 51-86

[32] S.S.K. Kwok et al.
A study of the importance of occupancy to building cooling load in prediction by intelligent approach
Energy Convers Manag
(2011)

[33] Z. Li et al.
Methods for benchmarking building energy consumption against its past or intended performance: an overview
Appl Energy
(2014)

[34] A novel hybrid model based on modal decomposition and error correction for building energy consumption prediction
2024, Energy

[35] Exploring the effects of a warmer climate on power and energy demand in multi-family buildings in a Nordic climate
2024, Environmental Advances

[36] V.G.R.C. Govindaraju, C.F. Tang
The dynamic links between CO2 emissions, economic growth and coal consumption in China and India
Appl Energy, 104 (2013), pp. 310-318