Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Prediction Models For Free Space Optical Systems: A Comparative Study Of Machine Learning Methods

Priteshkumar Bhikhubhai Vasava

Department of Electronics and Communication Engineering,
Government Engineering College, Bharuch, Gujarat, India.
Email: priteshvasava7187@gmail.com

The provided paper compares prediction models for forecasting signal quality in Free Space Optics (FSO) systems. The models are based on weather quality, turbulence levels, and environmental factors like the Air Quality Index (AQI). The models evaluated include Random Forest (RF), Gradient Boosting (GB), Long Short-Term Memory (LSTM), and a hybrid approach combining ensemble techniques and LSTM. The models' performance is assessed using metrics like Mean Squared Error (MSE) and R-squared. According to the data, the hybrid model outperforms the individual models, suggesting improved predictive power for FSO applications.

Index Terms—Atmospheric Turbulence, Free-Space Optical Communication, Machine Learning, Wireless Optical Communication, Attenuation.

I. INTRODUCTION

The Free Space Optics (FSO) is one such [11] modern communication technology based on the transmission of very high-speed data through the air, with modulated light beams serving as carriers. Unlike conventional optical fiber communications, which depend heavily on physical cables, FSO sets up a link between two points in free space, normally by means of laser sources, spanning several kilometers. This offers great versatility in installation [4,8,10] and use in an urban setting, other difficult locations, or any temporary installation, where laying of physical cables is not so economically feasible or logistic. FSO is an eyeful technology that greatly reduces the overhead of creation and maintenance of expensive infrastructure while providing great opportunities for installation in environments where temporary installations or rapid changes in installation are most required.

A. Positive features of FSO

High bandwidth FSO systems can be deployed to deliver input data at rates ranging from hundreds of megabits to several gigabits per second to support high-bandwidth applications like video conferencing, telemedicine, and cloud computing.

Ease of installation: FSO systems can be installed and operational just hours after a site is surveyed, often requiring little ground preparations. Because of this, they can be put into service without digging long trenches or major civil engineering projects.
Cost-Effectiveness: FSO technology does away with the need to build immobile infrastructure accompanying a high maintenance bill. It thus becomes economically viable to set up temporary or rapidly changing installations.
The environmental capabilities where FSO can operate efficiently include rugged terrains and urban landscapes in which conventional wired networks might have challenges.

B. Challenges in FSO
This presents a range of areas in which on FSO has begun to create serious challenges in the performance and reliability of modem systems on various accounts- for one based on weather conditions.
Circulation currents[3-7]: Since circulation currents will tend to ‘distort’ the beam by means of aerodynamic refraction when there are temperature variations causing non uniform depth distributions, some fading, that is a reduction of the signal amplitude below some threshold will occur. Such turbulence may, in fact, likely give rise to degraded signals especially for longer distances.

Atmosphere absorption: The fog and humidity [12] are some of the atmospheric conditions that can absorb the light beam leading to reduced beam intensity at the receiver. Such absorption seems to be more predominant in the presence of yet additional weather disturbances such as rain, snow, and fog.

Scattering: A scattering effect causes the light beam to be scattered by air suspensions such as dust, smoke and aerosol which result in a decrease transmitter signal at the receiver, hence increasing the noise level at the receiver. The concentration of particles in a unit volume of air causes the scattering to range widely on different levels.

C. Signal Predictive Modeling [12]

This predictive modelling could play a bigger role in optimizing the performance of free space optics in the presence of environmental factors. Engineers can utilize such mathematical formulas to aid in decision making towards enhancing the reliability and effectiveness of the system through the use of mathematical models that relate the environmental factors [17,21] and the signal quality.
1. Factors affecting quality signal prediction: The functions of predictive models are broadly defined in this area. These models estimate in advance the impact of certain atmospheric factors on the communication signal quality. The system operator is thus placed in a position to caution against such eventualities before they arise. The CRT foresight approach to this aspect helps to reduce communication system downtime, and improves management of the entire operational network, because in this case the CRT will be able to take such corrective actions before receiving a signal atrophy fault.

2. Designing Adaptive Systems: These enable an FSO system to adjust its transmission

characteristics in accordance with the current weather conditions without human intervention. For instance, if the system detects turbulence, it may switch to a more robust modulation technique that can withstand greater turbulence.
3. Total human capacity Planning: Using predictive modelling, the maximum possible capacity of the envisaged FSO link is calculated for various environmental conditions. Therefore, administrators may consider such maximum usage moments during their internal management of these periods. 4. Research and Development
The ongoing research on the predictive modelling of the performance of FSO remains relevant to new technology. It aids in the exploration of new materials, transmission techniques, and system design to counteract the negative influence of the environmental factors.
With general machine learning development will come lots of tricks in modelling complicated relations in data. In our work, we will describe various model capabilities of different machine learning fields, mainly concentrating on LSTM, RF, and GB, predicting signal quality based on environmental parameters.

II. BACKGROUND STUDY

When The Applications in urban networking, satellite communication, and disaster recovery are potential for the use of In Free Space Optical (FSO) communication systems due to their high-speed and cost-effectiveness. The interference that occurs during the transmission of data tends to cause degradation in signal reception due to environmental issues that include fog, rain, atmospheric disturbances, and scattering.

ML training is a very powerful option when weather factors are complex to explain, and it can cope with signal degradation in the future, as time-series predictions can be trained on large amounts of historical data.

Same goes with SVM and LSTM[1] -tree regression, which also falls under ML, has the capability to bridge the gap when predicting signal strength and quality in different weather conditions. LSTM being a neural network[11-12] excels at remembering sequential data due to its memory cell architecture. However, in this research, the strengths and weaknesses of each of these techniques are analyzed, especially in terms of their accuracy and speed of computation and practicality in the real world.

III. 3.PROBLEM FORMULATION

In addition to atmospheric conditions, turbulence and leakage are principal atmospheric factors causing enormous overall performance issues in open-space optical (FSO) conversation structures Conventional techniques often fail to correctly are expecting signal exceptional at pressure large eventualities. This assessment makes a specialty of exploring and trying out gadgets gaining knowledge of (ML) fashions that could predict the overall performance of FSO alerts the usage of regression and classification techniques together with LSTM. The main goal is to expand fashions that can provide correct, efficient, and real-time forecasts and growth system reliability. The look at may also compare model performance against environmental databases and recommend excellent practices for enhancing the power of the FSO communiqué gadget scattering [5].

IV. Research Methodology

A. Data Collection

The Data was collected from different sources, including environmental monitoring stations. The dataset contains features such as PM2.5, NO, NO2, CO, SO2, O3, AQI, Absorption, Turbulence, Scattering and weather quality.

B. Feature Engineering

To improve model performance, we have built additional features ofthe dataset.

- Normalized Features: Scaling the input features so that they have the same units.
- Lagged Features: Including previous time steps of data to capture any temporal dependencies.

C. Model Development

Various machine learning models which are helpful in Free space optical communication are discussed below.

1. Long Short-Term Memory (LSTM)

One Long Short-Term Memory (LSTM)[1-6] networks are a specialized type of Recurrent Neural Network (RNN) designed to overcome the limitations of traditional RNNs in capturing long-range dependencies within sequential data. LSTMs are particularly effective for tasks involving time-series predictions, natural language processing, and speech recognition, where contextual information from previous time steps is crucial for making accurate predictions.

Background on RNNs

Traditional RNNs process sequences of data by maintaining a hidden state that captures information from previous inputs. However, they suffer from the vanishing gradient problem, where gradients used for updating weights during training become too small, preventing the network from learning long-term dependencies. This makes it challenging for RNNs to retain relevant information over long sequences.

Architecture of LSTM

The LSTM architecture [1-6] introduces a unique cell structure designed to maintain and update the hidden state over time, allowing it to remember or forget information as needed. The key components of an LSTM unit include:

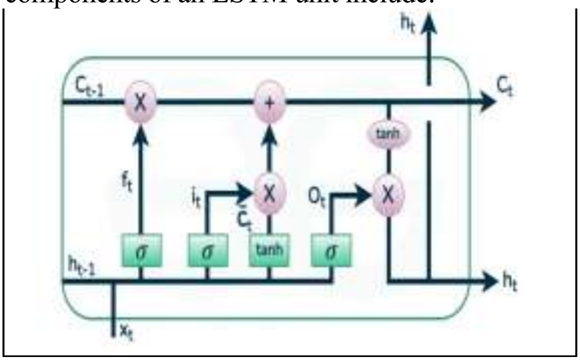

Fig. 1. Pictorial presentation of LSTM Algorithm

Cell State (Ct): This is the memory of the LSTM unit that carries information throughout the sequence. It is up to date via various gates.

Input Gate (it): This gate controls the volume to which new facts is added to the cellular kingdom. It uses the sigmoid activation function to output values among 0 and 1, determining how a great deal of the brand-new statistics to hold.

Forget Gate (toes): This gate decides what statistics to discard from the mobile state. Similar to the input gate, it outputs values between zero and 1, indicating how a whole lot of the preceding mobile state should be retained.

Output Gate (ot): This gate determines the output of the LSTM unit. It decides how plenty of the cellular state must be surpassed to the following hidden nation and in the end to the output of the LSTM.

Mathematical Formulation:

The operation of an LSTM unit can be described using the following equations:

Forget Gate:

The forget gate eliminates data that is no longer needed in the cell state. The gate receives two inputs, xt (the input at that specific moment) and ht-1 (the output of the previous cell), which are multiplied by weight matrices before bias is added. After being run through an activation function, the output is binary. When the output for a certain cell state is 0, the information is lost, and when the output is 1, it is saved for later use. The forget gate's equation is as follows:

$$ft=\sigma(Wf\cdot[ht-1,xt]+bf) \quad (1)$$

Where:

W_f represents the weight matrix associated with the forget gate.

[h_t-1, x_t] denotes the concatenation of the current input and the previous hidden state.

b_f is the bias with forget gate.

σ is the sigmoid activation function.

Input Gate:

The input gate modifies [1] the cell state by adding pertinent information. Using inputs ht-1 and xt, the sigmoid function is first used to regulate the information before filtering the values to be remembered in a manner akin to a forget gate. Next, a vector containing all possible values from ht-1 and xt is constructed using the tanh function, which produces an output ranging from -1 to +1. Finally, the useful information is obtained by multiplying the vector values by the regulated values. The input gate's equation is as follows:

$$it=\sigma(Wi\cdot[ht-1,xt]+bi) \quad (2)$$

$$C\text{^}t=\tanh(Wc\cdot[ht-1,xt]+bc) \quad (3)$$

We multiply the previous state by ft, disregarding the information we had previously chosen to ignore.

This represents the updated candidate values, adjusted for the amount that we chose to update each state value.

$$Ct=ft\odot Ct-1+it\odot C\text{^}tCt=ft\odot Ct-1+it\odot C\text{^}t \quad (4)$$

Where

⊙ denotes element-wise multiplication

tanh is tanh activation function

Output Gate:

The output gate's job is to extract valuable information from the current cell state so that it can be shown as output. The tanh function is first used to the cell to build a vector. Subsequently, the data is filtered by values to be retained using inputs ht−1 and xt, and the information is regulated using the sigmoid function. In order to send the values of the vector and the controlled values as an output and input to the following cell, they are finally multiplied. The output gate's equation is as follows:

$$ot=\sigma(Wo\cdot[ht-1,xt]+bo) \qquad (5)$$

2. Random Forest (RF)

The Random Forest [4-6]Procedure is a robust tree-based learning technique widely used in Calculator learning. During its education stage it constructs many conclusion trees. Each tree is formed using a random subset of the Information set which allows for the evaluation of a random selection of Characteristics at each split. This factor of haphazardness ensures that the apiece tree diagram is alone which helps to denigrate over fitting and improves general anticipation truth.

For making predictions the Procedure either averages the results for regression tasks or takes a majority vote for classification tasks. This mass decision-making work bolstered away Understandings from different trees results inch coherent and right outcomes. Random forests are frequently employed for both regression and classification tasks due to their ability to manage Complicated Information mitigate over fitting and provide dependable predictions across various situations.

Ensemble learning Representations can be compared to a group of friends with different skills coming together to solve a problem. They run care amp different squad of experts practical collaboratively to get decisions. Picture it as a team effort where friends with various expertise's add their strengths.

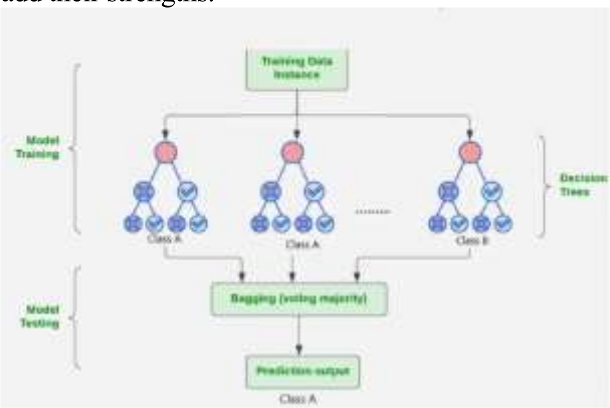

Fig. 2. Pictorial presentation of Random Forest Algorithm

Bagging and Boosting

Bagging is a group learning technique in which specific parts of the training data are used to train weekly models. For the regression problem the average forecast of the weekly observations is used to generate each subgroup model with replacement, and for the classification problem the majority vote is calculated [5-6]

Gradually boosting train models with more bases. In this process, each model tries to eliminate the errors introduced by the models that came before it. Each model is trained on a transformed dataset, with higher weights assigned to models misclassified by previous models. Weighted votes are used to predict victory.

Random Forest Processing Algorithm:

Step 1: Randomly select K data points from the training set.

Step 2: Create decision trees associated with selected data points (subsets).

Step 3: Select the number N for the decision trees you want to create.

Step 4: Repeat steps 1 and 2.

Step 5: Establish predictions for each decision tree for the new data points, then distribute them to the category that includes the majority.

Random Forest Operation [7-8]

The random forest algorithm works in several ways, which are described below:

Army Decision Trees: a random forest is an army of decision trees using the power of group learning. Each of these trees represents a specific expert focusing on a specific area of the data. Importantly, they behave differently, reducing the possibility that single-tailed particles will unduly influence the sample.

Random feature selection: Random Forest uses random feature selection to ensure that each decision tree in the cluster contributes to a unique perspective. Each tree has a few choices when training. Because each tree points to a different portion of the data due to this randomization, there are predictions for the entire group

Bootstrap aggregation or bagging: A key component of the random forest training method, bagging is the process of generating a large number of bootstrap samples from the original data set to create replacement samples with a different subset of data found in each decision Tree.

Decision and voting: Each decision tree in a random forest has a vote on future predictions. The final prediction in the classification task is determined by an all-tree approach or multiple predictors. The average of the predictions for each tree is calculated in the regression functions. This internal electoral system creates a collective and impartial decision-making process.

Salient Features of Random Forest

High prediction accuracy: The random forest works like a witch's circle in decision-making. Each diviner (decision tree) analyzes a different aspect of the story and combines their insights to create a powerful predictive mosaic. When two magicians work together, they can often create a more accurate picture than either.

Resistance to overloading: A random forest gradually guides trainees (decision trees) by an experienced trainer.

Each student acquires more knowledge than he can memorize every aspect of his training. This approach reduces the Chance of over fitting the model by preventing overuse during the training process.

Dealing with big data: Are you juggling a ton of data? A random forest planted nearby by a team of experienced

Decisive helpers will be nomadic. As each contributor takes a portion of the data set, the campaign is not only Incredibly fast but also accurate.

Presentation: Picture a random forest as a crime scene investigator, identifying the most important features or clues.

It explores the importance of ruling out signs and symptoms and helps focus on key factors affecting prognosis.

Integrated cross-validation: Random Forest works like a personal trainer to help you stay on track. Each of them also allocates hidden (bagged) information for testing when training the decision tree. This integrated certification ensures that your model performs well in addition to excellent training.

Managing unavailability: Just as datasets with unavailable values are part of life, so are uncertainties.

3. Gradient Boosting (GB) [3-6]

Gradient Boosting is a powerful boosting algorithm that combines several weak learners into strong learners, in which each new model is trained to minimize the loss function such as mean squared error or cross-entropy of the previous model using gradient descent. In each iteration, the algorithm computes the gradient of the loss function with respect to the predictions of the current ensemble and then trains a new weak model to minimize this gradient. The predictions of the new model are then added to the ensemble, and the process is repeated until a stopping criterion is met.

In contrast to AdaBoost, the weights of the training instances are not tweaked, instead, each predictor is trained using the residual errors of the predecessor as labels. There is a technique called the Gradient Boosted Trees whose base learner is CART (Classification and Regression Trees). The diagram below explains how gradient-boosted trees are trained for regression problems.

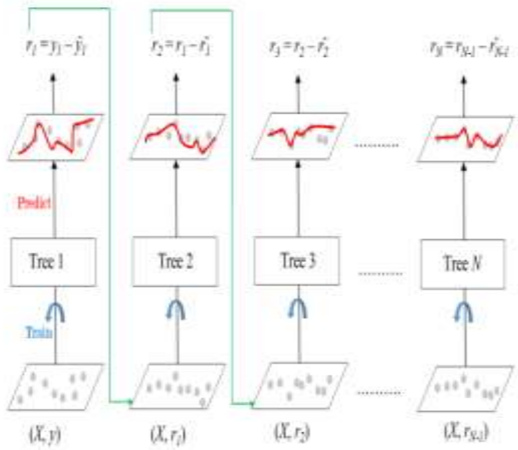

Fig. 3. Gradient Boosted Trees for Regression

The group comprises M trees. The labels y and the feature matrix X are used to educate Tree1. The schooling sets residual errors, r1, are located the use of the predictions categorized y1 (hat). After that, Tree2 is educated with the labels from Tree1's residual mistakes, r1, and the characteristic matrix X. The residual r2 is then ascertained using the expected consequences, r1 (hat). Until all M trees in the ensemble are educated, the process is repeated. This technique uses an essential parameter known as shrinkage. The term "shrinkage" describes how each ensemble tree's forecast shrinks after being multiplied by the learning rate (eta), which has a range of 0 to 1. The number of estimators and eta have a trade-off; in order to achieve a certain level of model performance, a decrease in learning rate must be made up for by an increase in estimators. Predictions are possible now that every tree has been taught. Every tree makes a label prediction, and the formula yields the final prediction.

y(pred) = y1 + (eta * r1) + (eta * r2) + ... + (eta * rN) (6)

4. Hybrid Model

This model combines the LSTM and RF/GB to leverage both sequential and ensemble modeling strengths, providing a more comprehensive prediction.[21]

Introduction to Hybrid Models

A hybrid model minimizes the drawbacks of every character version whilst maximizing its benefits with the aid of combining or extra device gaining knowledge of strategies. The aim is to growth prediction accuracy and generalizability with the aid of using clustering strategies like random woodland or gradient boom and deep gaining knowledge of fashions like LSTM. When dealing with complex, terrific statistics, while no unmarried method plays efficiently, hybrid fashions carry out specially well. Not in any respect.

Time-series analysis (LSTM) ensemble techniques are blended with hybrid models to seize the sign pleasant in open-space optics (FSO), wherein environmental factors like litter, absorption, and scattering can affect the signal first-class. This allows for the computation of complicated time dependencies and nonlinear fashions, which may be beneficial.

Key Components of the Hybrid Model

A hybrid model can integrate the following algorithms:

Long Short-Term Memory (LSTM): Captures temporal dependencies in time-series information by way of gaining knowledge of sequential styles.

Random Forest (RF): An ensemble technique that handles nonlinear relationships and decreases variance thru multiple selection timber.

Gradient Boosting (GB): Another ensemble technique that minimizes bias via studying from sequentially advanced weak rookies.

Working of the Hybrid Model

The role of hybrid representation is by combining predictions from multiple representations such as random forest, LSTM, and gradient boosting through nested weighted average or cascade methods. • Stacking: In cascading. Each performance is individually predicted. And these predictions are used as features in the metadata representation. (Usually a simple regression) to produce the result. • Weighted average: The predictions for each presentation are weighed according to their effectiveness. and use the weighted sum as the result. • Cascading: This involves feeding predictions from one agent to another as input. For the case of lstm yield, how to check the characteristics of arsenic random wood or slope improvement.

Mathematical Framework of the Hybrid Model

1. Individual Model Predictions:

Let LSTM, RF, and GB be the predictions from LSTM, Random Forest, and Gradient Boosting, respectively.

2. Weighted Average Approach for Hybrid Prediction:

$$\text{Hybrid}=w1\cdot \text{LSTM}+w2\cdot \text{RF}+w3\cdot \text{GB} \qquad (7)$$

Where w1, w2, w3are the weights assigned to each model, with w1+w2+w3=1.

1. Stacking Approach:

First, individual models make their predictions: XMeta= LSTM, RF, GB] (8)

A meta-learner (e.g., linear regression) takes XMeta as input and generates the final prediction Hybrid.

Advantages of the hybrid model

Durability: Using the advantages of different models. Many models will reduce the dangers of over- and under-assembly.

Increased accuracy: Hybrid models often perform better than single models. Because of its superior ability to capture complex patterns

Handles temporal and non-linear interactions: Hybrid models can more efficiently summarize unseen data because LSTM handles dependencies sequentially. And the cluster model handles non-linear interactions...

Applications in Free Space Optics (FSO)

The LSTM component in FSO communication systems helps to predict signal degradation caused by time-dependent components such as turbulence. On the other hand, random forest and gradient boosting are dependent factors. Non-linear environmental aspects such as diffusion and absorption. Hybrid models provide more accurate and reliable predictions of signal quality under changing atmospheric conditions than projections of individual models combined.

D. Model Evaluation

Models were evaluated using Mean Squared Error (MSE) and R^2 Score metrics:

Mean Squared Error (MSE):

$$MSE = (1/n) * \Sigma (y_i - \hat{y}_i)^2 \quad (9)$$

Where:

n = Number of data points

y_i = Actual value

$\hat{y}_i$ = Predicted value

R^2 Score

$$R^2 = 1 - (\Sigma (y_i - \hat{y}_i)^2 / \Sigma (y_i - \bar{y})^2) \quad (10)$$

Where:

y_i = Actual value

$\hat{y}_i$ = Predicted value

$\bar{y}$ = Mean of actual values

Evaluation of Hybrid Model

Hybrid model is a combination of different models (LSTM, Random Forest, and Gradient Boosting) which can be used to increase predictive performance. The purpose of it is each model outperforms the others by minimizing the MSE and maximizing the R2 score.

Equation:

$$\hat{y}\_hybrid = \alpha * \hat{y}\_LSTM + \beta * \hat{y}\_RF + \gamma * \hat{y}\_GB \quad (11)$$

Where:

$\hat{y}$_hybrid = Predicted value from the hybrid model

$\hat{y}$_LSTM, $\hat{y}$_RF, $\hat{y}$_GB = Predictions from LSTM, Random Forest, and Gradient Boosting models

- α, β, γ = Weights assigned to each model's prediction

This is a One-way ANOVA series sample which compares the prediction performances of the LSTM, Random Forest, Gradient Boosting, and the Hybrid Model. The performance of each model is measured with MSE and R^2 Score. The main aim is to show the Hybrid Model's superiority which features a higher accuracy of the predictions and outperforms all other models.

V. RESULTS AND DISCUSSION

A. Model Performance Comparison

The results obtained from the models are summarized in Table 1.

TABLE I Model Performance Comparison

Model	Mean Squared Error (MSE)	R^2 Score
LSTM	147842.35	-0.5973
Random Forest	13382.72	0.8554
Gradient Boosting	15299.20	0.8347
Hybrid Model	2723.97	0.9706

B. Observations from the Results

1. **LSTM Model:**

The LSTM model shows relatively poor performance with a high MSE value of 147842.35 and a negative R^2 score of -0.5973.
This indicates that the LSTM failed to capture the underlying data patterns effectively in this case, possibly due to hyper parameter settings or the complex nature of the dataset.

2. **Random Forest Model:**

The Random Forest model achieves a significant reduction in MSE to 13382.72 and an R^2 score of 0.8554, indicating it captures the data trends much better than LSTM.
This performance improvement is due to the ensemble nature of Random Forest, which reduces variance through averaging multiple decision trees.

3. **Gradient Boosting Model:**

With an MSE of 15299.20 and an R^2 score of 0.8347, Gradient Boosting performs similarly to Random Forest. However, it slightly underperforms in terms of both MSE and R^2.
This result highlights that Gradient Boosting may not be as effective on this dataset due to overfitting or the inability to generalize certain trends.

4. **Hybrid Model Performance:**

The Hybrid Model demonstrates superior performance, achieving the lowest MSE of 2723.97 and the highest R^2 score of 0.9706.
This indicates that the Hybrid Model effectively combines the strengths of individual models to capture complex data patterns.
The weights used in the Hybrid Model ensure optimal performance by balancing the contributions from LSTM, Random Forest, and Gradient Boosting.

C. Graphical Representation of Results

Figures 4 and 5 below illustrate the predicted vs. actual values for each model and a comparison of the models' performance metrics.

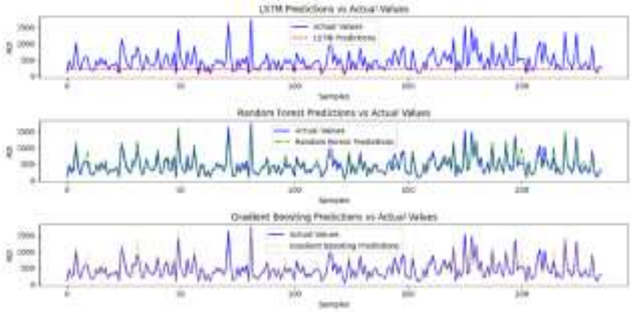

Fig. 4. Predicted vs. Actual Values for LSTM, Random Forest, Gradient Boosting, and Hybrid Model.

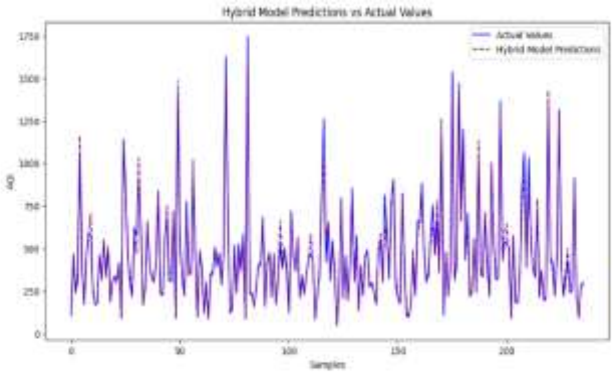

Fig. 5. Hybrid Model Predictions vs Actual Values

These visualizations confirm the Hybrid Model's superior performance over the individual models, as its predictions align more closely with the actual values, and it achieves the lowest error metrics.

D. Discussion of Results

The results indicate that the Hybrid Model outperforms the other models due to its ability to leverage the unique Strengths of LSTM, Random Forest, and Gradient Boosting. While LSTM struggled with the dataset, it was likely due to the long dependencies it tries to model, the Hybrid Model mitigates these challenges by incorporating efficiency of ensemble learning from Random Forest and Gradient Boosting.

- Random Forest and Gradient Boosting excel in capturing non-linear relationships in the data, while the LSTM model is known for handling sequential dependencies.
- By combining these techniques, the Hybrid Model achieves better generalization and robust predictive performance.

E. Conclusion of Results

The analysis shows that the Hybrid Model is the most effective approach for predicting environmental impacts on Free Space Optics (FSO) communication quality. With an MSE of 2723.97 and an R^2 Score of 0.9706, the Hybrid Model provides a reliable prediction mechanism, outperforming individual models. Future work can focus on fine-tuning the Hybrid Model further or incorporating additional environmental parameters to improve prediction accuracy.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

This research demonstrates the effectiveness of predictive modeling in enhancing the performance of Free Space Optics (FSO) communication by accounting for environmental factors such as atmospheric turbulence, absorption, and scattering. Four models—LSTM, Random Forest, Gradient Boosting, and a Hybrid Model—were implemented and compared to assess their ability to predict environmental impacts on signal quality.

The LSTM model, though well-suited for sequential data, struggled with complex environmental dependencies, yielding a high MSE and poor R^2 score. In contrast, Random Forest and Gradient Boosting captured non-linear relationships more effectively, producing reasonable prediction accuracy. However, the Hybrid Model—a combination of LSTM, Random Forest, and Gradient Boosting—outperformed all individual models by leveraging their complementary strengths.

With an MSE of 2723.97 and an R^2 score of 0.9706, the Hybrid Model provides the most reliable predictions, demonstrating that a combined approach is superior for optimizing FSO communication under varying environmental conditions. The success of the Hybrid Model confirms that ensemble-basedsolutions are more resilient and capable of generalizing across a broader range of conditions compared to individual models.

B. Future Work

Several avenues for future research are identified to further enhance the predictive capabilities of the proposed model:

1. Incorporation of Additional Environmental Factors:

Expanding the dataset to include other environmental parameters, such as rainfall, fog density, and temperature, could further improve the model's accuracy.

2. Real-Time Data Integration:

Implementing the model in real-time systems for non-stop monitoring of FSO links would possibly offer actionable insights for dynamic optimization.

3. Hyper parameter Optimization Techniques:

Future paintings could explore superior optimization strategies, along with Bayesian optimization or genetic algorithms, to extremely good tune the model parameters for higher universal performance.

4. Transfer Learning to Improve Adaptability:

Looking into transfer learning methods could help the model adapt better to new settings or data without needing lots of retraining.

5. Use in Internet of Things Applications: The model can be expanded to work in Internet of Things apps with mixed networks boosting how well it works across spread-out FSO networks.

6.Hybrid model modifications:

Further modifications of the hybrid model with the useful feature of evaluating weight variation for a single factor (LSTM, Random Forest, Gradient Boosting) should certainly result in further performance improvements overall in.
By analyzing those indicators, fate analysis can further help to build a robust and scalable FSO network, ensuring reliable statistical delivery in environmental conditions if even in the most difficult

ACKNOWLEDGMENT

We would like to thank everyone who is directly or indirectly associated withthis work. We Sincerely acknowledge the guidance provided from everyone.

REFERENCES

[1] Hochreiter, S., &Schmidhuber, J. (1997). "Long Short-Term Memory." Neural Computation, 9(8), 1735-1780.
[2] Chen, J., &Guestrin, C. (2016). "XGBoost: A Scalable Tree Boosting System." In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).
[3] Breiman, L. (2001). "Random Forests." Machine Learning, 45(1), 5-32.
[4] Friedman, J. H. (2001). "Greedy Function Approximation: A Gradient Boosting Machine." Annals of Statistics, 29(5), 1189-1232.
[5] Hochreiter, S., &Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[6] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50(1), 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0
[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. https://www.deeplearningbook.org/
[8]Muchisha, N., et al. (2023). Comparative analysis of LSTM, ARIMA, and hybrid models for time-series forecasting. Journal of Applied Forecasting, 56(2), 114-126.
[9] Wu, Z., et al. (2019). Deep learning models for short-term traffic flow prediction: A comparative study. Neural Computing & Applications, 31(8), 3125-3137. https://doi.org/10.1007/s00521-019-04209-x
[10] G. P. P. de Lima, C. M. F. de Almeida, and M. M. de Melo, "Machine learning approaches for predicting free space optical communication performance," Applied Optics, vol. 57, no. 1, pp. 125-134, 2018.
[11] M. A. Khalighi and O. Bouchet, "Free space optics: Principles and applications," IEEE Communications Surveys & Tutorials, vol. 16, no. 2, pp. 968-992, 2014.
[12] D. R. Kauffman and S. A. B. E. Mohd Ali, "Performance prediction of free-space optical communication systems using machine learning," Optics Express, vol. 27, no. 14, pp. 20040-20055, 2019.
[13] H. Elhoussainy, M. Elshafiey, and N. A. Ali, "Forecasting performance of free space optical communication systems with machine learning techniques," IEEE Access, vol. 9, pp. 115015-115026, 2021.
[14] D. Q. Wu, J. Li, and Y. Q. Wu, "A survey on achine learning techniques for free space optical communication," IEEE Access, vol. 7, pp. 147672-147683, 2019.
[15] B. K. Gupta, S. Sharma, and K. Singh, "Prediction of optical wireless communication performance using hybrid machine learning model," Journal of Optical Communications and Networking, vol. 11, no. 5, pp. 214-222, 2019.

[16] A. H. Mahmoud and M. I. Younis, "Predictive modeling for free space optics using deep learning techniques," Optics Communications, vol. 486, pp. 126-135, 2021.
[17] M. H. A. Rahman and J. Ali, "Machine learning for predicting FSO link performance under adverse weather conditions," Optical Engineering, vol. 59, no. 3, pp. 035103, 2020.
[18] L. N. T. Pham, A. T. T. Trinh, and T. T. H. Nguyen, "Machine learning for predicting free space optical communication link performance," Journal of Telecommunications and Information Technology, no. 3, pp. 60-67, 2020.
[19] Z. Zhao, X. Zhang, and H. Zhang, "Deep learning for predicting FSO performance under varying weather conditions," Journal of Optical Communications and Networking, vol. 13, no. 8, pp. 584-592, 2021.
[20] A. G. Al-Bawardy and A. R. T. P. K. R. B. Abd El-Monsef, "Performance prediction of free space optical communication systems using ensemble learning techniques," IEEE Access, vol. 8, pp. 136415-136423, 2020.
[21] B. Mukherjee, S. Roy, and D. Ghosh, "Adaptive predictive modeling for free space optics performance enhancement," Journal of Lightwave Technology, vol. 39, no. 2, pp. 299-305, 2021.

Dr. Priteshkumar Bhikhubhai Vasava
Born in Vyara, India, on January 7, 1987.
Educational Background: Bachelor of Engineering in Electronics and Communication, [VNSGU], 2004.Master of Engineering in Electronics and Communication, [GTU]2012. Ph.D. in Electronics and Communication, [GTU], 2024. Current Position: Assistant Professor, Electronics and Communication Department, Government Engineering College, Bharuch. Publications and Achievements: Published 3 research papers in reputed international journals and conferences. Research Interests: Artificial Intelligence (AI), Machine Learning (ML),Computer Networks Contact Information: Email: priteshvasava7187@gmail.