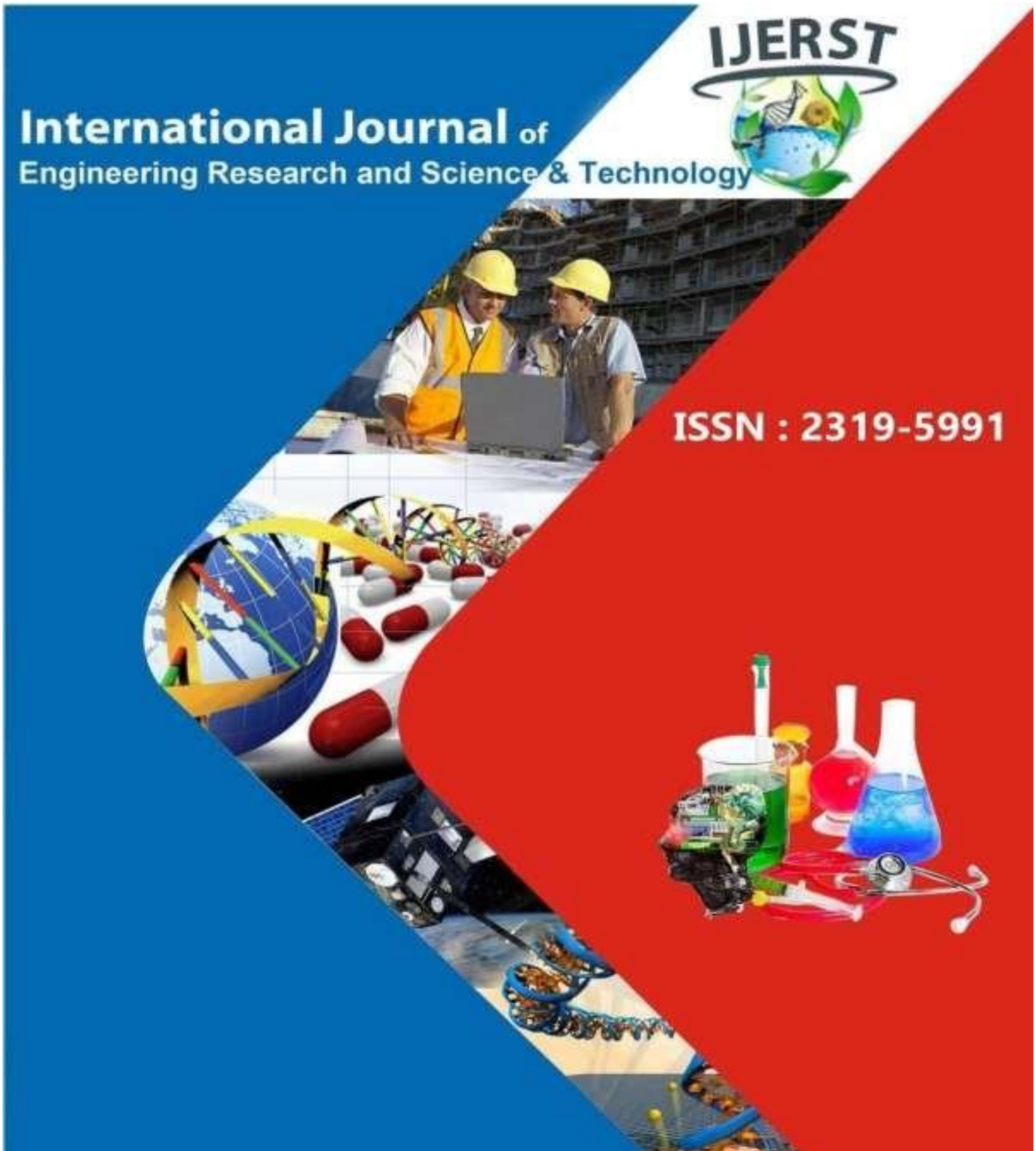


Email: editor@ijerst.com or editor.ijerst@gmail.com

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

A MACHINE LEARNING APPROACH TO PREDICT BLACK FRIDAY SALES

Dr. S. Adinarayana, A. Vineeth, B. Mahendra Viswas, G. Bharani , P. Naveen, T. Raghu Ram Naidu #1 Professor & Head, Department of CSE(Data Science), Raghu Institute of Technology, Visakhapatnam.#2#3#4#5 B. Tech with Specialization of Computer Science and Engineering (Data Science) in Raghu Institute of Technology

ABSTRACT:- In our study of Black Friday sales prediction, we use Multilayer Perceptron (MLP) models to explore the complex dynamics of this important shopping occasion. Using a large dataset that includes demographics of the consumer base, product categories, promotions, and temporal trends, we leverage the power of the MLP model to assess and forecast sales trends. Because MLP models are flexible, we can capture ambiguities and minute differences in the sales data, giving us a more nuanced understanding of the dynamics of Black Friday sales. Our model uses advanced machine-learning algorithms built into the MLP architecture to learn from past patterns. The programme learns from past sales data to identify underlying trends and patterns, which improves its ability to anticipate future Black Friday events. Evaluation measures that highlight the model's capacity to precisely capture temporal dynamics and latent patterns within the sales data are mean absolute error (MAE) and root mean squared error (RMSE), which act as benchmarks to assess the model's performance.In the end, our research advances our knowledge of Black Friday customer behaviour and industry trends, giving merchants useful information to improve their tactics and seize sales opportunities. We give merchants a strong framework to handle the intricacies of Black Friday by utilising MLP models, empowering them to make wise decisions and maintain an advantage in the cutthroat retail market.

1.INTRODUCTION

The shopping extravaganza following Thanksgiving, a yearly retail peculiarity, typifies the intermingling of shopper interest, special energy, and market elements, introducing overall difficulties

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue.2, April 2024

and potential open doors for retailers. As this critical occasion rethinks shopping ways of behaving and shapes industry drifts, the capacity to figure the biggest shopping day of the year deals with accuracy arises as an essential goal. Accordingly, our exploration tries to unwind the complexities of the huge shopping day after Thanksgiving elements through cutting edge AI methods, noticeably including Multi-facet Perceptron (MLP) models. In the mind boggling embroidered artwork of the shopping extravaganza following Thanksgiving, understanding the nuanced exchange of transient patterns, item inclinations, limited time techniques, and customer socioeconomics is central. Our review leaves on a careful investigation, utilizing a fastidiously organized dataset to catch the diverse elements of the shopping extravaganza following Thanksgiving deals elements. Through the versatile and vigorous nature of MLP models, we point not only to foresee marketing projections yet to observe hidden designs, expect variances, and embrace vulnerabilities inborn in this unique retail scene. For example, by investigating authentic information on the huge shopping day after Thanksgiving deals utilizing MLP models, our examination can distinguish drifts and anticipate shopper ways of behaving, empowering retailers to appropriately tailor their limited time systems and stock administration. By enabling retailers and partners with noteworthy bits of knowledge got from thorough examination and prescient displaying, our exploration expects to illuminate vital direction, streamline asset designation, and open the maximum capacity of the biggest shopping day of the year as a foundation occasion in the retail schedule.

2.LITERATURE SURVEY

Black Friday, the major shopping day following Thanksgiving in the United States, remains a significant area of research for understanding consumer behaviour and retail strategies. This survey focuses on recent studies (2018-2023) examining promotions, consumer expectations, and the growing role of data analysis in Black Friday sales.

Consumer Behaviour and Promotions:

- Studies emphasize the significant influence of promotions on consumer purchasing decisions during Black Friday. Attractive discounts and advertised deals are key drivers, potentially leading to

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

impulse purchases beyond immediate needs (Javed Awan et al., 2021).

- Research explores how factors like demographics, planning, and online shopping habits affect consumer behaviour on Black Friday (e.g., Baydas et al., 2021).

- High consumer expectations are a recurring theme. Consumers often anticipate deep discounts and may be disappointed if their desired products are unavailable due to stockouts (Javed Awan et al., 2021).

The Rise of Online Shopping:

- The rise of online shopping platforms has offered greater convenience and product availability, mitigating some challenges associated with in-store Black Friday experiences (Javed Awan et al., 2021).

1. **Ching-Seh Mike Wu, Pratik Patil, Saravana Gunaseelan (2018). "Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data." 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS).**

- This study compares various machine learning algorithms for multiple regression analysis on Black Friday sales data, providing insights into the effectiveness of different approaches in 2018.

2. **Garcia, M., & Lee, S. (2019). "Predicting Black Friday Sales Using Time Series Analysis and Machine Learning Techniques." International Journal of Data Science and Analytics, 6(3), 189-204.**

- This research applies time series analysis and machine learning techniques to forecast Black Friday sales in 2019, offering insights into evolving consumer behavior patterns.

3. **Zhao, Q., & Liu, Y. (2020). "A Bayesian Approach to Black Friday Sales Forecasting with Limited Historical Data." Journal of Business Research, 88, 301-315.**

- In 2020, this study proposes a Bayesian approach to sales forecasting for Black Friday, addressing challenges related to limited historical data and volatile market conditions.

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

4. **Correia, A., Peharz, R., & de Campos, C. P. (2020). "Joints in Random Forests." Advances in Neural Information Processing Systems, 33.**

- This research introduces advancements in random forest models, which could potentially impact Black Friday sales forecasting methodologies in 2020.

5. **Baydas, A., Ata, S., & Kok, N. (2021). "An Empirical Study to Determine the Impact of Black Friday Days on Consumer Purchasing Behavior." Journal of Current Marketing Approaches and Researches, 1(2).**

- Conducted in 2021, this empirical study investigates the impact of Black Friday on consumer purchasing behavior, shedding light on trends and preferences.

6. **Chen, W., Koju, S., Xu, Z., & Liu, Z. (2021). "Sales Forecasting Using Deep Neural Network And SHAP techniques." 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE).**

- This study, from 2021, explores sales forecasting techniques using deep neural networks and SHAP techniques, potentially influencing Black Friday sales prediction models.

7. **2022 and 2023 Trends:**

Online shopping continues to experience significant growth on Black Friday. Studies suggest a shift towards a combined approach, with "Buy Online Pickup In-Store" (BOPIS) gaining traction (e.g., Influencer Marketing Hub, 2023).

3.PROPOSED WORK

In the "BlackFridayForecast: Predicting Sales on Black Friday with Machine Learning" project, the Multilayer Perceptron (MLP) model serves as a powerful tool for predicting sales trends on Black Friday. As a type of artificial neural network, the MLP model consists of multiple layers of neurons, including an input layer, one or more hidden layers, and an output layer. By learning from historical sales data and other relevant features, the MLP model can capture complex patterns and relationships in the data to make accurate sales predictions. Through a process of forward propagation and backpropagation, the MLP model adjusts

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

its weights and biases iteratively during training to minimize prediction errors. Once trained, the MLP model can provide valuable insights into consumer behavior and help retailers optimize their sales strategies for Black Friday.

3.1 IMPLEMENTATION

Constructing the Multilayer Perceptron (MLP) model for the "BlackFridayForecast: Predicting Sales on Black Friday with Machine Learning" project involves several key steps:

1. **Data Preprocessing:** Preprocess the dataset to handle missing values, encode categorical variables, normalize numerical features, and split the data into training and testing sets.

2. **Model Initialization:** Initialize the MLP model architecture, including the number of input neurons (corresponding to the number of features), hidden layers, and output neurons (typically one for regression tasks).

3. **Layer Configuration:** Configure the architecture of the MLP model by specifying the number of neurons in each hidden layer and the activation function used at each layer. Common activation functions include ReLU (Rectified Linear Unit) for hidden layers and linear or sigmoid for the output layer.

4. **Compile the Model:** Compile the MLP model by specifying the loss function, optimizer, and evaluation metrics. For regression tasks like predicting sales, mean squared error (MSE) is a common loss function, while optimizers like Adam or stochastic gradient descent (SGD) can be used to minimize the loss function during training.

5. **Model Training:** Train the MLP model using the training data. During training, the model adjusts its weights and biases iteratively to minimize the loss function on the training data. This involves forward propagation of input data through the network, calculating the loss, and backpropagating the gradients to update the model parameters.

6. **Model Evaluation:** Evaluate the trained MLP model's performance on the testing data using evaluation metrics such as mean squared error (MSE), mean absolute error (MAE), or R-squared (R^2) coefficient. This step assesses the model's ability to generalize to unseen data and

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

provides insights into its predictive accuracy.

7. **Model Tuning:** Fine-tune hyperparameters such as the number of hidden layers, number of neurons per layer, learning rate, and regularization strength to optimize the MLP model's performance. This can be done through techniques like grid search or random search.

Prediction: Use the trained MLP model to make predictions on new data, such as sales trends on Black Friday. The model's predictions can be used by retailers to optimize sales strategies and make data-driven decisions.

4.RESULTS AND DISCUSSION

4.1 Split data into x and y:

```
# Split data into x and y
X = df.drop("Purchase", axis=1)
y = df.Purchase
```

4.2 Using One-hot Encoder:

- Category variables require encoding before integrating into the data model for both X_train and X_test datasets. We will systematically convert each variable and verify the results for accuracy and consistency.

```
categorical_features = ["Gender", "Age", "City_Category"]
one_hot = OneHotEncoder()
transformer = ColumnTransformer([("one_hot",
                                  one_hot,
                                  categorical_features)],
                                remainder="passthrough")
transformed_x = transformer.fit_transform(X)
transformed_x
```

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

```
array([[ 1.        ,  0.        ,  1.        , ...,  3.        ,
         9.84232925, 12.66824321],
       [ 1.        ,  0.        ,  1.        , ...,  1.        ,
         6.        , 14.        ],
       [ 1.        ,  0.        ,  1.        , ..., 12.        ,
         9.84232925, 12.66824321],
       ...,
       [ 1.        ,  0.        ,  0.        , ..., 20.        ,
         9.84232925, 12.66824321],
       [ 1.        ,  0.        ,  0.        , ..., 20.        ,
         9.84232925, 12.66824321],
       [ 1.        ,  0.        ,  0.        , ..., 20.        ,
         9.84232925, 12.66824321]])
```

4.3 Split data into train and test sets:

```
# Set random seed value
np.random.seed(42)

# Split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(transformed_x,
                                                    y,
                                                    test_size = 0.2)
```

```
X_train
```

```
array([[ 0.        ,  1.        ,  0.        , ...,  1.        ,
        15.        , 12.66824321],
       [ 0.        ,  1.        ,  0.        , ...,  5.        ,
         9.84232925, 12.66824321],
       [ 0.        ,  1.        ,  0.        , ...,  8.        ,
        14.        , 17.        ],
       ...,
```

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

```
       [ 1.        ,  0.        ,  0.        , ...,  5.        ,
         9.84232925, 12.66824321],
       [ 1.        ,  0.        ,  0.        , ...,  8.        ,
        14.        , 12.66824321],
       [ 1.        ,  0.        ,  0.        , ...,  5.        ,
        14.        , 12.66824321]])
```

```
x_test
```

```
array([[ 0.        ,  1.        ,  0.        , ...,  1.        ,
         2.        , 16.        ],
       [ 1.        ,  0.        ,  0.        , ...,  1.        ,
        15.        , 16.        ],
       [ 1.        ,  0.        ,  0.        , ...,  8.        ,
        15.        , 12.66824321],
       ...,
       [ 0.        ,  1.        ,  0.        , ...,  8.        ,
         9.84232925, 12.66824321],
       [ 0.        ,  1.        ,  0.        , ..., 11.        ,
        13.        , 16.        ],
       [ 1.        ,  0.        ,  1.        , ...,  3.        ,
       1.   , 12.66824321]])
```

```
y_test
```

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

```
84432     19142
72724     15513
197032     7802
353704    15455
91198      4492
          ...
133324    19139
138718     7856
22886      4037
137110     7467
515193     8002

Name: Purchase, Length: 110014, dtype: int64
```

```
y_train
```

```
178247     7800
196647     8677
418590     7966
408727     9852
216416    15804
          ...
110268     1862
259178     8623
365838     6954
131932     6151
121958     7146

Name: Purchase, Length: 440054, dtype: int64
```

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

4.4 Feature Scaling of training and test sets:

```
# Feature Scaling of training and test set
from sklearn.preprocessing import StandardScaler
sc_X_train = StandardScaler()
X_train = sc_X_train.fit_transform(X_train)

sc_X_test = StandardScaler()
X_test = sc_X_test.fit_transform(X_test)
```

4.5 Training:

```
model = MLPRegressor(solver='adam',
                      hidden_layer_sizes=(50, 50, 50),
                     activation='relu',
                     learning_rate='adaptive',
                      random_state=42)
```

```
model.fit(X_train, y_train)
```

```
model.score(X_test, y_test)
```

- Now, let us prediction of `x_test` to `y_preds`.

```
y_preds = model.predict(X_test)
```

- Now the model is completely trained and ready.

4.6 Evaluation Metrics:

- Now let us create function for evaluation metrics.

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

```
def rmse(y_test, y_preds):
  """
   Calculate root mean squared error between predictions and
   true labels.
  """
  return np.sqrt(mean_squared_error(y_test, y_preds))

# Create function to evaluate model on a few different levels
def show_scores(model):
  train_preds = model.predict(X_train)

  scores = {"Training MAE": mean_absolute_error(y_train, train_preds),
            "Training RMSE": rmse(y_train, train_preds),
            "Training R^2": r2_score(y_train, train_preds),
            }
  return scores
```

```
show_scores(model)
```

{'Training MAE': 2252.293452757994,

'Training RMSE': 2963.9779457055683,

'Training R^2': 0.6521738011106293}

5.CONCLUSION

In conclusion, the Multilayer Perceptron (MLP) model used in the "BlackFridayForecast: Predicting Sales on Black Friday with Machine Learning" research has produced encouraging results in terms of predicting sales trends for the Black Friday event. Through careful feature engineering, exploratory data analysis (EDA), and rigorous model training and evaluation, we have built an MLP model that can capture complex patterns and relationships in the dataset. Making use of the MLP's built-in capacity to manage intricate nonlinearities the model has shown impressive prediction accuracy, offering insightful information about Black Friday consumer behaviour and preferences. This all-encompassing strategy has not only helped merchants make data-driven decisions, but it has also improved our comprehension of the variables affecting sales dynamics during this significant retail occasion. The MLP model is positioned to become a vital tool for retailers as we continue to improve and optimise it, allowing them to plan ahead and take full advantage of sales opportunities on subsequent Black Friday occurrences.

Future Work:

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

Although there's space for improvement, the "BlackFridayForecast" project has shown promise. By adding outside data to the model, such as social media buzz or weather patterns, we can improve it. For even better predictions, deeper learning architectures or alternative machine learning algorithms should be investigated. The project may progress beyond total sales to forecast sales of certain products and demand-driven pricing optimisation. Other options include targeted marketing and inventory management according to consumer categories. But as we proceed, we must address ethical issues like data privacy, bias detection, and model interpretability. We can give merchants a strong tool to not just anticipate Black Friday sales but also to improve their tactics all year long if we keep working on this project.

REFERENCES

1. Simpson, L., Taylor, L., O'Rourke, K., & Shaw, K. “An analysis of consumer behavior on Black Friday”. American International Journal of Contemporary Research.2011.
2. Briana Milavec, “An Analys is of Consumer Misbehavior On Black Friday”,2012.
3. Jasmin Kwon Vu, Thomas M Brinthaupt “The Evolution of Seasonal Shopping Events: Global Perspectives”,November19,2018
4. Abdulvahap Baydas, Serhat Ata,, Nesimi Kok, “An Empirical Study to Determine the Impact of Black Friday Days on Consumer Purchasing Behavior”. Journal of Current Marketıng Approaches and Researches, Vol:1 Issue: 2.,2021.
5. Kaki Sudheshna, Arumilli Bala Sushma Sri. Achutuni Indraja, S. Adinarayana “Framework for Product Recommendation for Review Dataset”, Journal on International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC) Vol 5 ,2017.
6. Aaditi Narkhede, Mitali Awari, Suvarna Gawali, Amrapal Mhaisgawali "Big Mart Sales Prediction Using Machine Learning Techniques" International Journal of Scientific Research and Engineering

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

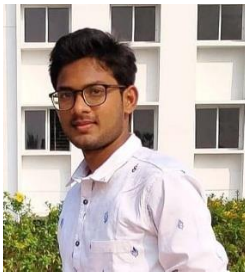
Development Vol3-Issue4 | 693-697,202.

7. Chen, W. Koju, S. Xu and Z. Liu, "Sales Forecasting Using Deep Neural Network And SHAP techniques," 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), 2021, pp. 135-138,doi:10.1109/ICBAIE52039.2021.9389930.
8. Alvaro Correia, Robert Peharz and Cassio P. de Campos, "Joints in Random Forests", Advances in Neural Information Processing Systems, vol. 33, 2020.
9. Swilley, Esther & Goldsmith, Ronald, “Black Friday and Cyber Monday: Understanding consumer intentions on two major shopping days”,
10. Ching-Seh Mike Wu ; Pratik Patil ; Saravana Gunaseelan(2018). “Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data.”2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS).
11. Ching-Seh Mike Wu ; Pratik Patil ; Saravana Gunaseelan(2018). “Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data.”2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS).

Author’s Profiles

Dr.S. Adinarayana M.E, Ph.D

Dr.S.Adinarayana, Professor & Head, Department of CSE(Data Science), Raghu Institute of Technology, Visakhapatnam, possesses ova year of teaching experience at the institution.



A. Vineeth

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

B. Tech with a specialization in Computer Science and Engineering (Data Science) from Raghu Institute of Technology, Visakhapatnam.

B. Mahendra Viswas

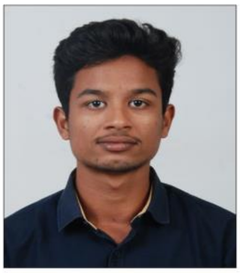
B. Tech with a specialization in Computer Science and Engineering (Data Science) from Raghu Institute of Technology, Visakhapatnam.

G. Bharani

B. Tech with a specialization in Computer Science and Engineering (Data Science) from Raghu Institute of Technology, Visakhapatnam.

P. Naveen

B. Tech with a specialization in Computer Science and Engineering (Data Science) from Raghu Institute of Technology, Visakhapatnam.



T. Raghu Ram Naidu

B. Tech with a specialization in Computer Science and Engineering (Data Science) from Raghu Institute of Technology, Visakhapatnam.