FOOD DELIVERY TIME PREDICTION USING CNN- LSTM

Jyothi K [1], Raksha H L[2], Raksha S[3], Rakshitha G Upadhya[4], Sowmyashree C B[5]

Department of ISE JNN College of Engineering

---***---

Abstract - **To tackle the challenge of predicting food delivery times, the study introduces a novel approach crucial for streamlining food delivery service operations. The proposed model combines CNN (Convolution Neural Network)s and (LSTM) Long Short-Term Memory networks, leveraging the unique strengths of each design. The CNN component analyzes delivery data trends on a daily and weekly basis, identifying geographical patterns such as rush hours and seasonal variations. Meanwhile, the LSTM component focuses on recognizing spatiotemporal relationships and retaining data over time to effectively model sequential data. Compared to standard procedure, the model's simultaneous incorporation of spatial and temporal information leads to more precise predictions. This enhanced accuracy holds proven strategy for food delivery services, enabling them to more accurately estimate delivery times, reduce delays, and ultimately improve overall customer satisfaction.**

Keywords: (LSTM) Long short term memory, Time prediction, CNN (Convolution Neural Networks).

1.INTRODUCTION

As urbanization advances, traffic congestion emerges as a prominent issue, significantly disrupting daily routines. Despite the availability of live traffic updates through navigation apps and websites, accurately estimating travel times remains challenging. The existing real-time traffic updates may not always accurately represent the conditions upon reaching the destination emphasizing the need for more precise travel time estimation based on traffic speed prediction. Although traditional statistical and machine learning techniques have been utilized for forecasting traffic velocity, recent progressions in deep learning particularly CNN(Convolutional Neural Network)s and LSTM (Long Short-Term Memory) models, hold promise in addressing this challenge. These advanced techniques leverage extensive traffic data and powerful computational resources to improve prediction accuracy. Specifically, CNN-LSTM models have shown effectiveness in predicting traffic conditions by extracting space and time features. However, current research predominantly focuses on traffic speed prediction, overlooking other critical applications like food delivery time estimation. Accurately predicting food delivery times is essential for optimizing delivery operations and enhancing customer satisfaction, requiring precise estimations of travel times considering various factors such as road conditions, traffic congestion and delivery routes. In this study, we introduce a CNN-LSTM model customized specifically for food delivery time forecasting. aiming to address the unique challenges associated with estimating delivery times. By harnessing CNN's ability to extract temporal features related to daily and weekly traffic patterns and LSTM's capability to capture spatiotemporal dependencies, our model aims to provide accurate and reliable predictions for food delivery times. Through extensive simulations and validation using real- world delivery data, it showcase the efficacy of our suggested CNN-LSTM model in enhancing the precision and dependability of food delivery time estimation in contrast to conventional techniques. This research contributes to optimizing food delivery operations, ultimately leading to enhanced customer experiences and satisfaction. The proposed CNN-LSTM model in enhancing the correctness and completeness of food delivery time prediction compared to traditional methods. This research contributes to optimizing food delivery operations, ultimately leading to improved customer experiences and satisfaction.

The paper consists of Literature Survey followed by methodology in section 3. Along with this, the paper also consists of information about the CNN-LSTM model in section 4 and its results in section 5. The final part of the paper consists of conclusion and references.

2. LITERATURE SURVEY

The paper introduces a pioneering method to enhance the precision of predicting hand gesture timing by integrating signal processing and deep machine learning techniques, extracting crucial timing attributes from raw data [1]. Following this preprocessing step, the data is inputted into a CNN(Convolutional Neural Network) for classification, capitalizing on its effectiveness in analyze a series of abstract symbols to determine their meaning and context. Another contribution involves in suggesting a model that combines CNN(Convolutional Neural Network)s with a transformer architecture to predict the multivariate time series data [5]. This approach captures both temporal patterns and inter-variable relationships concurrently, leading to improved prediction accuracy in contrast to current methods. The paper proposes a novel deep learning framework called Multivariate (Convolutional Neural Network). (MVCNN), tailored for classifying multivariate time series data [6]. This architecture tackles obstacles in CNN application to time series by suggesting a tensor scheme to convert data into 3D tensors, facilitating the utilization of local interactions among variables.

The paper introduces a classification strategy founded on LSTM for forecasting time series trends [3]. This method utilizes LSTM (Long Short-Term Memory) structures to capture both long-term and short-term dependencies in sequential data, showcasing better performance compared to traditional auto-regression models. The study presents a fresh model, LSTM-attention-LSTM, forecast future events using past data trends[2]. This model utilizes two LSTM models along with an attention mechanism to enhance accuracy, particularly for non-stationary multivariate data. Experimental findings underscore its superiority over current approaches, investigating the control of the attention mechanism across various time steps.

3.METHODOLOGY

To analyze the delivery time of food we follow eight steps: first data collection Secondly pre-processing the data, data split, model selection and training the CNN-LSTM model. Finally evaluating the model to see its delivery time prediction illustrate in Fig. 1.

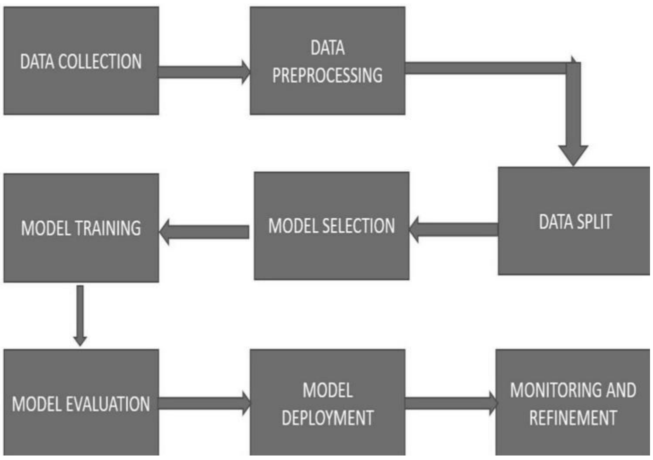

Fig-1: Block diagram of CNN-LSTM model.

3.1 Data Collection

Food delivery dataset consists of 11 attributes. It contains 45,594 entries. Each entry contains ID, Delivery person id, Delivery person Age, Delivery person Ratings, Restaurant latitude, Restaurant longitude, Delivery location latitude, Delivery location longitude, Type of order, Type of vehicle, Time taken(min), weather condition, Traffic condition, distance.

3.2 Preprocessing

Preprocessing stage ensues to validate its accuracy and suitability for analysis. This phase encompasses tasks such as data cleaning, managing missing values, outlier detection and removal, and formatting the information into an appropriate structure for analysis. Moreover, techniques like data normalization or scaling may be utilized to standardize features onto a comparable scale. In food delivery time prediction The Haversine formula is essential for calculating the geographic distance between the food preparation site and the delivery destination, utilizing latitude and longitude coordinates. This calculation forms the foundation for accurate estimations, allowing delivery services to address for various factors such as traffic conditions and optimal routes. By incorporating these variables, delivery times can be predicted more precisely, securing efficient and punctual order deliveries, thereby enhancing both customer gratification and operational effectiveness.

3.3 Data Split

The gathered and preprocessed data is commonly split into two subsets training set and test set. Training set is utilized to train the prediction model, and the other is employed to assess its performance. This partitioning of data is typically done randomly, guaranteeing that both subsets accurately represent the overall data distribution.

3.4 Model Selection

In choosing the right prediction model, the project's requirements and available data are carefully considered. Numerous machine learning algorithms are generating, including support vector egression, decision trees, linear regression, gradient boosting, random forests, and deep learning architectures like CNN (Convolutional Neural Network) or LSTM (Long Short-Term Memory) networks. This decision is guided by factors such as the dataset's characteristics, the intricacy of the task, and the expected performance standards.

3.5 Model Training

During this phase, the chosen prediction model undergoes training using the provided training data. The model assimilates information from input features and their corresponding delivery time labels to comprehend patterns and correlations. Through the training process, optimization methods are employed to iteratively fine-tune model parameters, aiming to minimize prediction errors.

3.6 Model Evaluation

Following training, the model undergoes performance evaluation using the test dataset. Assessment criteria like mean absolute error (MAE), mean squared error (MSE), or root mean squared error (RMSE) are often utilized to measure prediction precision. The model's performance is scrutinized to ascertain if additional enhancements or adjustments are necessary.

4. CNN –LSTM MODEL

CNN (Convolutional Neural Network) offer numerous advantages for tasks involving time prediction. Their primary strength lies in their ability to autonomously acquire

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

hierarchical features from unprocessed input data like audio, images, time-series. This hierarchical feature learning enables CNNs to effectively capture intricate patterns and interconnection present in the data, making them particularly adept at handling sequential or time-dependent information. Furthermore, CNNs exhibit robustness to variations in input information, for instance shifts, distortions, or noise, which proves vital for accurate time prediction in practical situations where data quality may be compromised. However, a notable drawback of CNNs in time prediction tasks is their computational complexity, especially with longer time series. This complexity can result in prolonged training and latency, especially when addressing large datasets or high-resolution temporal data, necessitating significant computing power and potentially hindering real-time applications.

LSTM (Long Short-Term Memory) networks present several advantages for time prediction endeavors. One notable benefit is their adeptness at capturing long-term dependencies within sequential data, rendering them well-suited for time series forecasting. This potential is facilitated by memory cells and gating mechanisms regulating information flow, enabling LSTMs to retain particulars over extended durations. Moreover, LSTMs exhibit proficiency in handling variable-length sequences, a crucial attribute for modeling temporal data (time series data) characterized by fluctuating sequence lengths. However, a drawback of employing LSTMs for time prediction tasks lies in their computational complexity and training duration especially when addressing extensive datasets or intricate architectures. Training LSTMs necessitates substantial computational resources and time investment, potentially constraining their applicability in real-time prediction scenarios where rapid processing is paramount.

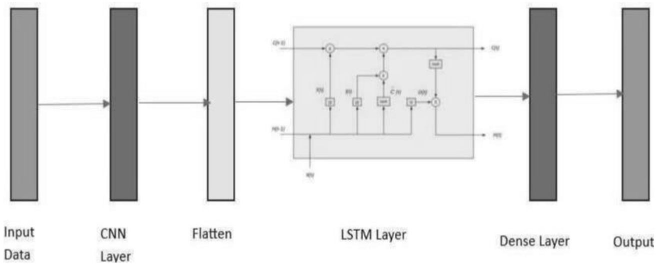

Fig-2: CNN-LSTM model.

A CNN-LSTM model integrates CNN (Convolutional Neural Network) with LSTM (Long Short-Term Memory) networks to capitalize distinct strengths of each architecture for time prediction applications. CNNs excel at capturing spatial patterns within data, making them well-suited extracting features from sequential data datasets like time series. By employing (Convolutional Neural Network) layers on the input sequence, the CNN can acquire hierarchical representations of temporal data, discerning critical patterns and features. Subsequently, the LSTM network processes sequential information over time. LSTMs are proficient at capturing long-term dependencies and temporal dynamics, rendering them consummate for tasks where data order and context are pivotal, such as time prediction. By amalgamating CNNs for feature extraction and LSTMs for sequential modeling, the CNN-LSTM model adeptly captures both spatial and time series patterns in the data, resulting in enhanced performance for time prediction tasks. This hybrid proceeding towards enabling the model to discern intricate relationships within time series data, culminating in more accurate predictions compared to employing either architecture in isolation.

The input data comprises the raw information processed by the model. In the realm of food delivery time prediction, this encompasses various factors such as order details (items, quantities), customer information (location, order history), time-related aspects (time of day), and external conditions (weather, traffic). Although not commonly utilized in this domain, CNN(Convolutional Neural Network) layers can be beneficial for analyzing spatial data, such as maps or geographical information. If incorporated, the CNN layer applies filters to extract spatial features. Subsequently, a Flatten layer reshapes the output into a one-dimensional array, facilitating the transition to recurrent layers like LSTM (Long Short-Term Memory). LSTM layers play pivotal part in capturing temporal dependencies within the sequential data, essential for precise delivery time prediction. Following the LSTM layers, dense layers contribute by capturing complex arrangements and relationships within the data. Ultimately, the result layer delivers predictions, typically employing a single neuron for regression tasks, yielding continuous estimates of delivery time.

5. RESULTS

Our proposed CNN-LSTM model was implemented and run in laptop with processor Intel(R) Core™ i7-3520 CPU, 8GB RAM and 64- bit Windows Operating System.

5.1 Experimental Setup

The developed food delivery time forecasting model using PySpark's machine learning library. The LSTM (Long short term memory) model was trained utilizing the Adam optimization algorithm, with a size of testing dataset of 0.2. The batch size was configured to 32 for a duration of 20 Epochs. The chosen number of epochs, set to 20, reflects a common starting point in neural network training methodologies. This selection facilitates iterative learning processes over the training data, enabling the model to capture intricate patterns inherent within the dataset. However, the determination of the optimal epoch count necessitates meticulous observation of both training and validation performance metrics to mitigate the risk of overfitting while maximizing computational efficiency.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Adjustments to the epoch count should be made judiciously, guided by discernible trends in the training and validation loss, with careful consideration afforded to the employment of early stopping mechanisms as a safeguard against overfitting phenomena. The Test loss graphs of the three models are depicted in Fig-5.

	ID	Delivery_person_ID	Delivery_person_Age	Delivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_location_longi
0	0x4607	INDORES13DEL02	37	4.9	22.745049	75.892471	22.765049	75.91
1	0xb379	BANGRES18DEL02	34	4.5	12.913041	77.683237	13.043041	77.81
2	0x5d6d	BANGRES19DEL01	23	4.4	12.914264	77.678400	12.924264	77.68
3	0x7a6a	COIMBRES13DEL02	38	4.7	11.003669	76.976494	11.053669	77.02
4	0x70a2	CHENRES12DEL01	32	4.6	12.972793	80.249982	13.012793	80.28
5	0x9bb4	HYDRES09DEL03	22	4.8	17.431668	78.408321	17.461668	78.43
6	0x95b4	RANCHIRES15DEL01	33	4.7	23.369746	85.339820	23.479746	85.44
7	0x9eb2	MYSRES15DEL02	35	4.6	12.352058	76.606650	12.482058	76.73
8	0x1102	HYDRES05DEL02	22	4.8	17.433809	78.386744	17.563809	78.51
9	0xcdcd	DEHRES17DEL01	36	4.2	30.327968	78.046106	30.397968	78.11

Fig-3: Different attributes list

rd	...	Weatherconditions	Road_traffic_density	Vehicle_condition	Type_of_order	Type_of_vehicle	multiple_deliveries	Festival	City	Time_taken(min)	distance
00	...	Sunny	High	2	Snack	motorcycle	0	No	Urban	24	2.050806
00	...	Stormy	Jam	2	Snack	scooter	1	No	Metropolitian	33	14.086096
00	...	Sandstorms	Low	0	Drinks	motorcycle	1	No	Urban	26	1.083801
00	...	Sunny	Medium	0	Buffet	motorcycle	1	No	Metropolitian	21	5.457067
00	...	Cloudy	High	1	Snack	scooter	1	No	Metropolitian	30	4.333926

Fig-4: Distance calculated using Haversian Formula

MSE = $1/n\Sigma_{i=1}^{\eta}(y_i - \hat{y}_i)^2$

n represents the total count of data points.

y_i is the actual target value.
$\hat{y}_i$ is the predicted target value.
Mean Squared Error (MSE) measures the average squared difference between actual and predicted values. It's a common metric used to evaluate regression models, including neural networks. Lower MSE indicates better model performance

Table-1: Comparison of models

Models	CNN	LSTM	CNN-LSTM
Test Loss	67.7108 s^2	68.121 s^2	66.2281 s^2

Table-1 compares the results of LSTM, CNN, CNN-LSTM test loss.

```
Epoch 1/20
912/912 [==============================] - 10s 7ms/step - loss: 159.3318 - val_loss: 76.6951
Epoch 2/20
912/912 [==============================] - 6s 6ms/step - loss: 71.8600 - val_loss: 70.3869
Epoch 3/20
912/912 [==============================] - 6s 6ms/step - loss: 69.4423 - val_loss: 68.2838
Epoch 4/20
912/912 [==============================] - 6s 6ms/step - loss: 68.5432 - val_loss: 67.6483
Epoch 5/20
912/912 [==============================] - 6s 7ms/step - loss: 68.2313 - val_loss: 67.4593
Epoch 6/20
912/912 [==============================] - 6s 7ms/step - loss: 67.8295 - val_loss: 67.2988
Epoch 7/20
912/912 [==============================] - 6s 7ms/step - loss: 67.5469 - val_loss: 66.8120
Epoch 8/20
912/912 [==============================] - 7s 7ms/step - loss: 67.4149 - val_loss: 67.1077
Epoch 9/20
912/912 [==============================] - 7s 8ms/step - loss: 67.2990 - val_loss: 66.7769
Epoch 10/20
912/912 [==============================] - 6s 7ms/step - loss: 67.2169 - val_loss: 67.1879
Epoch 11/20
912/912 [==============================] - 6s 6ms/step - loss: 67.1775 - val_loss: 66.8574
Epoch 12/20
912/912 [==============================] - 6s 6ms/step - loss: 67.0570 - val_loss: 66.5227
Epoch 13/20
912/912 [==============================] - 6s 6ms/step - loss: 66.9891 - val_loss: 66.5878
Epoch 14/20
912/912 [==============================] - 6s 6ms/step - loss: 66.8997 - val_loss: 66.4467
Epoch 15/20
912/912 [==============================] - 6s 6ms/step - loss: 66.8932 - val_loss: 66.6435
Epoch 16/20
912/912 [==============================] - 6s 6ms/step - loss: 66.7997 - val_loss: 66.4231
Epoch 17/20
912/912 [==============================] - 6s 6ms/step - loss: 66.8047 - val_loss: 66.2970
Epoch 18/20
912/912 [==============================] - 6s 6ms/step - loss: 66.7550 - val_loss: 66.4521
Epoch 19/20
912/912 [==============================] - 6s 7ms/step - loss: 66.6653 - val_loss: 65.9776
Epoch 20/20
912/912 [==============================] - 6s 7ms/step - loss: 66.6549 - val_loss: 66.0434
```

Fig-5 model fit of CNN-LSTM

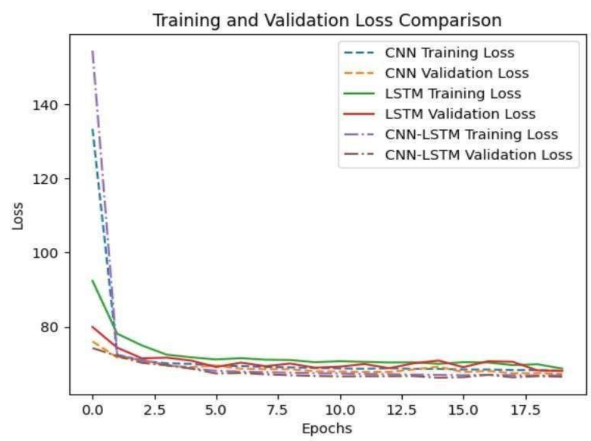

Fig-6. Training and Validation loss comparison on models

The details of the achieved Training and Validation Loss can be observed in Figure 5. By comparing the Test loss obtained by CNN model, LSTM model and the combined CNN-LSTM model, it is clear that the Test loss for combined CNN-LSTM model is less. Therefore, the used model is employed for time prediction.

6. CONCLUSION

This paper aimed to tackle the challenges encountered by food delivery services in accurately predicting delivery times for customers, thereby improving operational effectiveness and enhancing customer contentment. Through the utilization of advanced methods of Machine Learning and data analysis, a forecasting model was created utilizing a dataset. The combination of the CNN-LSTM model resulted in lower test loss, thus improving prediction accuracy. This model has the potential to significantly enhance the entire food delivery process by providing precise estimates of delivery times and facilitating efficient coordination and route optimization. Proposed work marks a significant advancement in optimizing food delivery operations, maximizing driver utilization, and improving overall operational efficiency within the food delivery industry.

REFERENCES

[1] G. Park, V. K. Chandrasegar and J. Koh, "Accuracy Enhancement of Hand Gesture Recognition Using CNN," in IEEE Access, vol. 11, pp. 26496-26501, 2023

[2] X. Wen and W. Li, "Time Series Prediction Based on LSTM-Attention-LSTM Model," in IEEE Access, vol. 11, pp. 48322-48331, 2023.

[3] Y. Liu, Z. Su, H. Li and Y. Zhang, "An LSTM based classification method for time series trend forecasting," 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), Xi'an, China, 2019, pp. 402-406.

[4] Irem Islek, Sule Gunduz Oguducu ,"Use of LSTM for Short-Term and Long-Term Travel Time Prediction" 2018 , CEUR Workshop Proceedings 2018, vol -2482

[5] D. -K. Kim and K. Kim, "A Convolution Neural Network Transformer Model for Multivariate Time Series Prediction," in IEEE Access, vol. 10, pp. 101319-101329, 2022

[6] C. -L. Liu, W. -H. Hsaio and Y. -C. Tu, "Time Series Classification With Multivariate Convolution Neural Network," in IEEE Transactions on Industrial Electronics, vol. 66, no. 6, pp. 4788-4797, June 2019.

[7] Z. Wang, X. Su and Z. Ding, "Long-Term Traffic Prediction Based on LSTM Encoder-Decoder Architecture," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 10, pp. 6561- 6571, Oct. 2021

[8] R. Ghanbari and K. Borna, "Multivariate Time-Series Prediction Using LSTM Neural Networks," 2021 26th International Computer Conference, Computer Society of +Iran(CSICC), Tehran, Iran, pp. 1-5, 2021.

[9] Preeti, R. Bala and R. P. Singh, "Financial and Non-Stationary Time Series Forecasting using LSTM Recurrent Neural Network for Short and Long Horizon," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-7

[10] C. -H. Yang, C. -H. Wu and C. -M. Hsieh, "Long Short-Term Memory Recurrent Neural Network for Tidal Level Forecasting," in IEEE Access, vol. 8, pp. 159389-159401, 2020.

[11] H. Yao, F. Wu, J. Ke, X. Tang, Y. Jia, S. Lu, P. Gong, J. Ye, and Z. Li, ''Deep multi-view spatial–temporal network for taxi demand prediction,'' in Proc. AAAI Conf. Artif. Intell, vol. 32, no. 1, LA, USA, 2018, pp. 1–8.

[12] A. Sherstinsky, ''Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network,'' Phys. D, Nonlinear Phenomena, vol. 404, no. 1, pp. 51–53, 2020.

[13] Y. Yu, X. Si, C. Hu, and J. Zhang, ''A review of recurrent neural networks: LSTM cells and network architectures,'' Neural Comput., vol. 31, no. 7, pp. 1235–1270, Jul. 2019.

[14] J. Gehring, M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin, "Convolutional sequence to sequence learning," arXiv preprint arXiv:1705.03122, 2017.

[15] J. Zhao, F. Deng, Y. Cai, and J. Chen, ''Long short-term memory–fully connected (LSTM-FC) neural network for PM2.5 concentration prediction,'' Chemosphere, vol. 220, pp. 486–492, Apr. 2019, doi: 10.1016/j. chemosphere.2018.12.128.