International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Novel Approach for Human Behaviour Prediction Using Deep Learning Algorithms

Dr. Babhuti Kashyap[1], Dr. Vrinda Sachdeva[2], Dr. Abhinandan Shirahatti[3], Dr. Gurwinder Singh[4], Amit Arya[5], Dr. Senthilkumar Jagatheesan[6], Chaman Kumar[7]

Submitted: 28/08/2023 **Revised**: 25/10/2023 **Accepted**: 03/11/2023

Abstract: Predicting human behaviour is a complex and multifaceted endeavour with implications spanning various domains, from healthcare and marketing to security and social sciences. This research paper delves into the application of deep learning techniques for the prediction of human behaviour. The study explores the use of neural networks, including Long Short-Term Memory (LSTM), convolutional neural networks (CNNs), and other advanced deep learning architectures in capturing intricate patterns and dependencies in human behaviour data.

We begin by discussing the importance of human behaviour prediction, its real-world applications, and the challenges associated with this task. We also highlight the significance of feature engineering and data preprocessing techniques in enhancing prediction accuracy. The research emphasizes the critical role of data quality, model interpretability, and ethical considerations in the deployment of deep learning for human behaviour prediction. Moreover, it addresses the ongoing research challenges and future directions in this field, such as addressing biases, handling sparse data, and integrating multimodal data sources. In conclusion, this paper underscores the promise of deep learning in advancing our ability to predict human behaviour, with the potential for transformative applications in numerous sectors. The findings presented herein contribute to the ongoing dialogue on harnessing artificial intelligence for a better understanding of and adaptability to human behaviour.

Keywords: *Human behaviour prediction, Deep learning, LSTM (Long Short-Term Memory), CNN (Convolutional Neural Network), Data preprocessing*

1. Introduction

Recognizing human actions in videos is a difficult but essential part of action recognition and video processing. The underpinning for activity recognition is constant monitoring of human behaviour. It has formed the foundation for several uses, including healthcare and geriatric monitoring, injury detection in sports, human position estimate, and home surveillance. Human activity detection from video sequences is still a challenging subject despite significant advancements in the field. The hardest aspect is definitely the feature extraction. It does affect the algorithm's performance, as it speeds up computations and simplifies logic. Human activity recognition [1] relies on manually generated local characteristics extracted from RGB footage captured by 2D cameras that are incapable of handling complicated human actions. Some strategies depend on the extraction of a moving individual from their backdrop. Human activity can also be characterized through the use of motion tracking. Human tracking is used to manage its motion and construct trajectories over the scene. Humans have an intuitive ability for tracking. Once the objects' velocities go too high, the problem becomes difficult to solve. Human activity recognition is the study of predicting the kinds of behaviours and objectives that one or more agents will engage in. Their past behaviour and activities will be analyzed to determine the best course of action. Common daily actions for a human being include walking, running, sitting, standing, lying down, ascending and descending stairs, and more. If HAR could be used to detect and analyze people's actions, it may lead to innovative solutions in the sectors of child and elder care combined with tools and methods from the Internet of Things [1]. In the case of a child left at a day care while their parents went to work, the HAR might be used to forecast the child's behaviour and let the parents know whether or not their child is in danger. Caregivers or

[1]*Assistant Professor Psychology Centre for online education, Department of Psychology, Chandigarh University, Mohali*
Email- babhuti.e13687@cumail.in
[2]*Associate professor ITS engineering college, greater Noida Computer science department Email- vrinda08@gmail.com*
[3]*Associate Professor Department of Computer Science and Engineering KIT College of Engineering (Autonomous) Kolhapur Maharashtra*
Email- abhinandans2010@gmail.com
[4]*Assistant Professor, Department of AIT-CSE, Chandigarh University, Punjab, India. Email- singh1001maths@gmail.com*
[5]*Assistant Professor Madanapalle Institute of Technology & Science*
Email- amitaryaaol@gmail.com
[6]*Associate Professor,Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,*
No.42, Avadi-Vel Tech Road, Vel Nagar, Avadi, Chennai, Tamil Nadu, India. Email- adhityaresearch@hotmail.com
[7] *department of computer science IIMT College of engineering greater Noida*
Email- chamankumar31@gmail.com

guardians might utilize this technology to keep their elderly charges safer by preventing them from engaging in risky behaviours. Therefore, HAR has the potential to deliver appealing answers to pressing human issues across a wide range of contexts, from personal fitness to gaming to security domains to the health care business and beyond. Recent advances in Deep Learning have made CNN and RNN architectures increasingly popular, but researchers have only begun to explore how these networks might be used to train time series of inertial sensor data [2, 3]. Models of deep learning include CNN and In order to develop discriminative features from raw sensor data, RNNs take a data-driven approach. External or wearable sensors, such as accelerometers and gyroscopes, are typically used to track human activity. Accelerometer data monitors the speed at which individuals move, while gyroscope data measures the angular velocity at which they move of the acts. [4]Therefore, it will be crucial to create enough automated systems to handle and evaluate the whole dataset that these sensors offer. Based on this, large amounts of raw data will be gathered, processed to create a feature vector, and then utilized as input to learning algorithms to produce an activity recognition model [5]. To get the highest possible identification accuracy, it is crucial to use a well-trained, well-organized prototype.

2. Literature Survey

In this, the author [1] presented a study on the prediction of human activities, both static and dynamic, using machine and deep learning models. They explored various algorithms and techniques to classify and predict human activities. The core objective was to improve the accuracy of activity recognition in real-world scenarios. The authors concluded that their deep learning models, combined with machine learning approaches, provided promising results for activity prediction, which can have applications in fields such as healthcare and human-computer interaction.

In this, the author [2] presented a novel approach to human behaviour prediction in smart home environments. They employed deep reinforcement learning techniques to predict and understand human actions within the context of a smart home. The core objective was to enhance the adaptability and automation of smart home systems by accurately predicting and responding to human behaviour. The authors' work aimed to improve the overall user experience and efficiency of smart home environments. Their approach demonstrated the potential of using deep reinforcement learning for human behaviour prediction in the context of IoT and home automation.

In this, the author [3] introduced an approach for pedestrian prediction using deep neural networks. Their core objective was to enhance pedestrian safety and autonomous driving by accurately predicting the behaviour of pedestrians in urban environments. The authors utilized deep neural networks for planning and predicting pedestrian movements, emphasizing the importance of real-time decision-making in autonomous vehicles. Their work demonstrated that deep learning techniques could contribute to the development of advanced driver assistance systems and autonomous vehicles, making them safer and more efficient in navigating complex urban scenarios.

In this the author [4] introduced a core objective was to develop a model that could predict an individual's personality traits based on their social media activity and content. The authors employed various machine learning algorithms to analyze text and behaviour on social media platforms and make predictions about a user's personality. The results of this work are valuable for personalized content recommendations, targeted advertising, and understanding online behaviour. The study highlighted the potential of machine learning in extracting insights from social media data for various applications, including personalized user experiences and content curation.

In this, the author [5] presented a core objective was to enhance traffic safety and management by accurately predicting areas with a higher risk of accidents in urban environments. The authors utilized deep learning techniques to predict accident-prone areas. This research contributes to improving traffic safety and optimizing resource allocation for accident prevention and emergency response. The study demonstrated the potential of deep learning for urban planning and intelligent transportation systems, with the aim of reducing accidents and enhancing overall traffic management.

In this, the author [6] focused their core objective was to improve the accuracy of predicting short-term crime incidents in urban areas by incorporating human mobility patterns. The authors leveraged deep learning techniques and analysed human mobility data to gain insights into how human movement influences crime rates. By combining mobility data and advanced deep learning models, they aimed to provide valuable insights for law enforcement and urban planning to better allocate resources and enhance safety measures. This research highlights the potential of integrating mobility data and deep learning for more effective crime prediction and prevention strategies in urban environments.

In this, the author [7] conducted a core objective was to evaluate and compare different deep learning models for the recognition of complex scenarios involving multiple activities and objects within visual data. By examining various deep learning architectures, the authors aimed to identify the most effective model for MAMO recognition.

In this, the author [8] explored the use of deep learning for time-to-lane-change prediction. The core objective was to develop a model that could accurately predict when a vehicle would change lanes in an intelligent transportation system context. By leveraging deep learning techniques, the authors aimed to enhance traffic safety and management by predicting lane change behaviour, which is critical for autonomous vehicles and advanced driver assistance systems. The research demonstrated the potential of deep learning in contributing to intelligent transportation systems by improving predictions related to lane changes, thereby enhancing traffic safety and overall traffic flow management.

In this, the author [9] Nielsen, focuses on practical time series analysis with an emphasis on prediction using statistics and machine learning. The core objective of the book is to provide a comprehensive and practical guide for analyzing time series data, making predictions, and applying statistical and machine learning techniques effectively. By covering a wide range of topics related to time series analysis, Nielsen's work aims to equip readers with the knowledge and skills needed to work with time series data in various fields, including finance, economics, and forecasting.

In this, the author [10] focused on core objective was to develop a model that could recognize and predict the daily activities of individuals based on images captured from a first-person perspective, such as through wearable cameras. By leveraging deep learning, the authors aimed to create a system that can assist in activity recognition for applications like healthcare, personal assistance, and augmented reality. This research showcased the potential of deep learning in understanding and predicting human behaviour based on visual data, thereby contributing to advancements in wearable technology and human-computer interaction.

In this, the author [11] introduced an unsupervised user behaviour prediction algorithm for smart homes, based on a combination of machine learning and neural network techniques. The core objective was to develop an approach that could predict user behaviour within a smart home environment without relying on explicit labels or supervision. By employing machine learning and neural network models, the authors aimed to enhance the adaptability and automation of smart home systems, ultimately improving the user experience and energy efficiency. This research demonstrated the potential of unsupervised learning in predicting and understanding user behaviour in smart home settings, offering advancements in home automation and energy conservation.

In this, the author [12] conducted an exploratory study on predicting depression using machine learning methods and smartphone behavioural markers. The core objective was to develop a model that could identify signs of depression based on patterns in smartphone usage and behaviour the authors aimed to provide a valuable tool for early detection and monitoring of depression by demonstrating the potential of using smartphone data for mental health assessment, ultimately leading to improved mental health care and early intervention for individuals at risk of depression.

In this, the author [13] presented a core objective was to assess and compare various methods and techniques used to predict human behaviour in the context of smart and intelligent environments. By analysing different approaches and their performance, the authors aimed to provide insights into the effectiveness of predictive models for applications like smart homes, healthcare, and automation in human behaviour prediction and helping researchers and practitioners choose the most suitable methods for specific use cases, thereby advancing the development of intelligent systems.

In this, the author [14] focused on the application of machine learning for the efficient assessment and prediction of human performance in collaborative learning environments. The core objective was to develop models and techniques that can assess and predict how well individuals perform in collaborative learning scenarios. By leveraging machine learning methods, the authors aimed to improve the assessment of collaborative learning outcomes, which is crucial in educational and homeland security contexts. Their research contributes to the advancement of personalized learning and security efforts by providing tools for ultimately enhancing decision-making and resource allocation.

In this, the author [15] focused on the core objective was to develop a system that can recognize and predict. By employing deep learning algorithms, the author aimed to enhance the accuracy and applicability of activity recognition and prediction in various domains, including healthcare, sports, and assistive technology. This research highlights the potential of wearable sensors and deep learning for understanding and forecasting human behaviour, offering advancements in personalized healthcare and human-computer interaction.

A. WISDM Dataset

For our work we have used WISDM data set. Time series sensor data from 51 test individuals completing 18 activities, each lasting 3 minutes, and sampled at 20Hz via a smartphone and wristwatch. There were 36 people that took part in the study. The UCI machine-learning repository offers this for download. There are 1,098,207 records in the collection, and they cover six different activities: walking, jogging, going up and down stairs, sitting, and standing. User, Activity, Time, X Acceleration, Y Acceleration, and Z Acceleration are the columns. This dataset was originally somewhat lopsided, with 38.6% of records involving walking, 31.2% involving jogging, 11.2% involving being upstairs, and 9.1% involving being seated and 4.4% involving being standing.

B. CNN Architecture

The section of the human brain known as the visual cortex served as an inspiration for the development of CNNs. A neuron in a particular layer will only have connections with a small fraction of the neurons in the layer below it, rather than being connected to all of the neurons in that layer. There is also a term for what are known as Convolutional Neural Networks, or CNNs. There is a divide between CNN and other media outlets. Input, output, convolutional, relu, pooling, and fully connected layers make up the traditional CNN architecture. These layers are referred to as the input layer, the output layer, the pooling layer, and the fully connected layer, respectively. Backpropagation is often utilised because it is the method that most effectively converges the estimation error while also ensuring that the final convolution is appropriately weighted. The goal of the convolutional operation is to extract the high-level features in an effort to offer a more vital and subtle link amid the classification process.

C. LSTM Architecture

RNNs with longer short-term memory (LSTM) are able to learn and recall across very long arrangements of input data. This ability allows them to do exceedingly complex tasks. As a consequence of this, one of the most common uses of LSTM is in the field of time series analysis. Activation functions are not used in LSTM's recurrent layers. The values that have been stored are not modified. The problem of a gradient that gradually decreases over time does not arise during the training of LSTM [5, 6].

LSTMs are often constructed in "blocks" or cells, and each block may include anywhere from three to four gates (Fig. 1), which may include input and output gates in addition to forget gates. LSTMs are typically created in "blocks" or cells. The problem of disappearing gradients can be solved with LSTM by applying the gating concept [6].

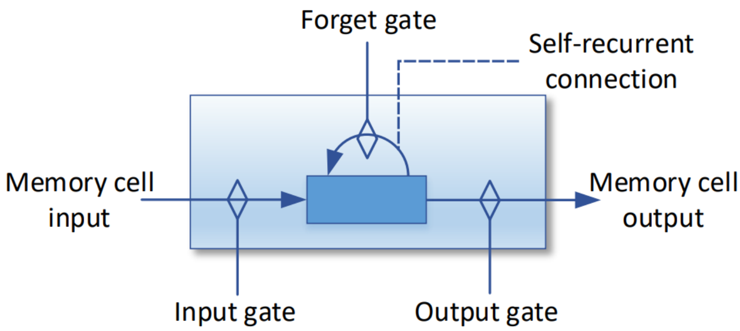

Fig. 1. Illustration of an LSTM memory cell [7].

Because LSTMs are able to train from data on unprocessed time series, they represent a feasible alternative for sequence classification that does not need the hand-engineering of input properties by subject matter experts. It is possible to input data from a number of sensors, such as accelerometers and gyroscopes, all at the same time into the model.

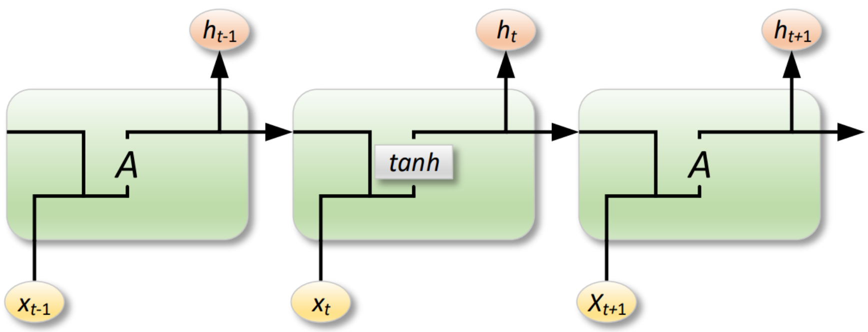

Fig. 2. Standard RNN module

Through exposure to sequences of observations and the subsequent extraction of features from those sequences, the model is provided with the knowledge necessary to map its internal properties to a wide variety of activity types.

D. Human Activity Recognition with CNN and LSTM

HAR has been widely a vast study field in the past decade. The use of HAR on information collected by smartphones and other sensors has advanced fast in recent years, and it now has a wide range of useful applications. Recent developments in the field of mobile sensing have made it possible for people to keep tabs on anything from their emotional state and commuting habits to their sleep and exercise habits [6]. Activity recognition issues have also been attacked with statistical learning techniques. We've utilised a variety of methods, including Naive Bayes and K-Nearest Neighbour (KNN), to identify seven distinct movements, including walking, running, leaping, and so on. Systems have gotten more heuristic [10], but their design has become increasingly complex and requires specialised knowledge.

Numerous implementations of HAR employing CNN and LSTM models have been shown to be successful in the literature. In [8], using the UCI-HAR dataset. They also demonstrate a method that incorporates convolutions into the LSTM units' input reading process. Overall, the CNN LSTM model has reached 90.6% accuracy, whereas the Convolutional LSTM model provided around 90% accuracy. In [9], we see a way for employing CNN LSTM, with the convolutional output serving as the LSTM input. Different time frames have been used for convolution, and a "Time Distributed layer" has been implemented for time series analysis. Additionally, using LSTM networks for modelling behaviour has led to the development of several probabilistic models for behaviour prediction [10].

3. Methodology

The methodology for predicting human behaviour by combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks begins with comprehensive data collection and preprocessing, ensuring data quality. Next, the CNN is utilized to extract spatial features, particularly valuable for visual and sensor data, aiding in the recognition of spatial aspects of human behaviour. Simultaneously, LSTM is integrated to model temporal dependencies, capturing how behaviour evolves over time. By combining the outputs of the CNN and LSTM, the model becomes adept at recognizing both spatial and temporal patterns, resulting in a powerful predictive tool. Training and evaluation on labeled data are crucial, alongside hyperparameter tuning and cross-validation for robustness. Deploying the model for real-time or batch prediction, interpretability, and addressing ethical considerations complete the methodology, contributing to applications such as healthcare, security, and smart environments.

The methodology for predicting human behaviour by combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks:

1. *Data Collection and Preprocessing:*

The initial step in the methodology involves the collection of a diverse dataset representing human behaviour. This dataset may encompass various sources, including sensor data, video footage, or other relevant data inputs. Ensuring data quality is pivotal, and this is achieved through thorough data preprocessing. Data cleaning, normalization, and organization tasks are performed to eliminate inconsistencies and noise, ensuring that the data is reliable and ready for analysis.

2. *Feature Extraction using CNN:*

Convolutional Neural Networks (CNNs) are employed for spatial feature extraction. They shine when it comes to data sources like images, videos, or sensor readings. The CNN layers, equipped with convolutional filters and pooling operations, enable the identification of critical patterns and features within the data. This makes the model proficient at recognizing spatial aspects of human behaviour, such as specific gestures, postures, or objects within a scene. In cases where the data is sequential, it might undergo a transformation process to become suitable for input into the CNN. This transformation effectively converts sequential data into image-like representations, enhancing the model's ability to capture spatial information.

3. *Temporal Modelling with LSTM:*

To understand the temporal evolution of human behaviour, Long Short-Term Memory (LSTM) networks are introduced. LSTMs are tailored for modelling sequences and are excellent at capturing how behaviour patterns change over time. They are particularly skilled at recognizing transitions in activities and predicting future actions based on historical behaviour data. This temporal modelling component is pivotal in understanding how and when different activities or behaviours occur.

4. *Model Integration:*

The methodology advances by combining the outputs of the CNN and LSTM networks within a unified architecture. This fusion may involve concatenating the feature vectors extracted by the CNN with the hidden states of the LSTM. The combined model now possesses the ability to consider both spatial and temporal information, allowing it to recognize and predict human behaviour more accurately and comprehensively.

5. *Interpretability and Visualization:*

Tools and techniques for interpreting the model's predictions are implemented. This is crucial for understanding the reasoning behind the model's decisions. Visualization and interpretation methods provide insights into the features and temporal aspects that influence human behaviour, enabling a more profound understanding of the model's predictions.

4. Expiremental Setup

Because of its widespread usage in research on intelligent settings and activity recognition, this dataset was selected for analysis. Our makes it easier to make comparisons between the work of our publication and that of other academics working in both domains. The dataset is the result of observing a man who was 26 years old and living in a three-room flat where there were 14 binary sensors installed. These sensors were installed in every accessible surface, including cupboards, refrigerators, freezers, and stalls. There were a total of 21,220 sensor events and 247 activity occurrences that were recorded during the course of the period of 28 days. The following activities were marked with their respective annotations: "LeaveHouse," "UseToilet," "TakeShower," "GoToBed," "Prepare Breakfast," "Prepare Dinner," and "Get Drink."

We conducted three types of studies to verify the architecture's efficacy:

- Architecture experiments: we tested alternative designs, altering the number of LSTMs and completely linked dense layers.
- The second set of tests examined the results of varying the duration of the sequence of actions that were entered.
- Experiments were conducted to determine the impact of including input action timestamps.

Experiments with the architecture (detailed in Table 1) included varying LSTM layer counts, LSTM layer types (normal and bidirectional [12]), fully connected layer counts, and fully connected layer sizes. To further investigate the benefits of embeddings, we compared them to the conventional method of employing one-hot vectors for representing the activities.

Table 1. Design innovations in buildings. Dropout-free regularisations are good. LSTM No. indicates LSTM layer size ina recurrent neural network. Dense No. is the number of ReLU-activated completely linked layers, and all designs have a softmax-activated final layer. ReLU-activated fully linked layers are compact. Measure the series length against the input action sequence. Activities are encoded using embeddings or one-hot vectors.

ID	Dropout	LSTM #	LSTM Size	Dense #	Dense Size	Sequence Length	Coding
A1	0.4	1 (Standard)	512	1	1024	5	Embedding
A2	0.8	1 (Standard)	512	1	1024	5	Embedding
A3	0.8	1 (Standard)	512	2	1024	5	Embedding
A4	0.8	2 (Standard)	512	2	1024	5	Embedding
A5	0.2	1 (Standard)	512	5	50	5	Embedding
A6	0.8	1 (Bidirectional)	512	2	1024	5	Embedding
A7	0.8	1 (Standard)	512	2	1024	5	One-hot vector
A8	0.8	1 (Standard)	512	1	1024	5	One-hot vector
A9	0.8	2 (Standard)	512	2	1024	5	One-hot vector

Experiments with different sequence lengths are summarised in Table 2 below. For these experiments, we used a network with a dropout regularisation of 0.8, one LSTM layer of size 512, two fully connected layers of size 1024 each with ReLU activation and a final fully connected layer of size 1024 with softmax activation, and varied only the sequence length of the input action. Experiment A3 also made advantage of this setup.

Table 2: Prolonged experiments in a sequence. The A3 setup was utilised for all experiments.

ID	Sequence Length
S1	3
S2	1
S3	4
S4	6
S5	10
S6	30

Finally, we conducted experiments using time to test out various strategies for factoring in the timing of events within the input sequence. We looked into three possibilities. For the first setup (T1), we combined the outputs of two parallel LSTM layers—one for the action embeddings and one for the timestamps—just before the fully linked layers (a late fusion method. The second setup (T2) uses an early fusion technique to combine the action embeddings and timestamps before feeding them into a single LSTM layer. The third setup (T3) included a slow fusion technique [15] where the embeddings were linked to an LSTM layer, whose output was combined with the timestamps and delivered to another LSTM layer. Experiment A3's dropout regularisation was employed across all configurations, as was an LSTM layer size of 512, two ReLU-activated fully connected layers of 1024 size, and a third fully connected layer with softmax activation.

5. Result

The outcomes of each experiment design are tabulated in Tables 3-5. More LSTM layers (A4) or more tiny, densely linked layers (A5) did not improve model accuracy in the architectural tests (Table 3), regardless of the number of predictions. Using LSTMs that can learn in both directions (A6) also had a negative impact. In most cases, the best outcomes were obtained with dropout regularisation values of A2 and A3. All things considered, the A3 setup provided the greatest outcomes.

As can be seen in the table, the results achieved by utilising embeddings were superior than those produced by using one-hot vectors in the identical configurations (A3 against A7, A2 versus A8, and A4 versus A9). The same architecture may simulate user behaviour more accurately when the actions are represented in a more semantically rich way. When compared to alternative representation schemes, such one-hot vectors, the benefit of employing embeddings is that the training procedure to produce them is fully unsupervised.

For the sequence-length trials (Table 4), the greatest outcomes were attained utilising sequences with a length of 4 (S3) or 5 (A3) actions. The average action duration of the activities in each deployment was a key factor in determining the ideal sequence length in this circumstance. This value should be changed in each unique scenario in order to attain the greatest outcomes, and it cannot be generalised.

comparable tasks that occur at different times (for example, making breakfast and preparing supper).

Table 3. Experiments in architectural design accuracies for varying sets of predictions

ID	acc_at_1	acc_at_2	acc_at_3	acc_at_4	acc_at_5
A1	0.4487	**0.6367**	0.7094	**0.7948**	0.8461
A2	0.4530	0.6239	**0.7222**	0.7692	0.8504
A3	**0.4744**	0.6282	0.7179	0.7905	**0.8589**
A4	0.4444	0.5940	0.6965	0.7735	0.8247
A5	0.4402	0.5982	0.7136	0.7820	0.8418
A6	0.4487	0.6068	0.7136	0.7905	0.8376
A7	0.4572	0.6153	0.7094	0.7820	0.8376
A8	0.4529	0.5811	0.7051	0.7735	0.8376
A9	0.4102	0.5940	0.7008	0.7777	0.8247

Table 4. Experiments with varying sequence lengths and their effects on prediction accuracy.

ID	acc_at_1	acc_at_2	acc_at_3	acc_at_4	acc_at_5
A3	**0.4744**	0.6282	0.7179	0.7905	**0.8589**
S1	0.4553	0.5957	0.7021	0.8	0.8553
S2	0.4255	0.6255	0.7021	**0.8085**	0.8382
S3	0.4658	**0.6452**	**0.7264**	0.7948	0.8504
S4	0.4700	0.6196	0.6965	0.7692	0.8461
S5	0.4592	0.6351	0.7210	0.7896	0.8369
S6	0.4192	0.5589	0.6593	0.7554	0.8122

Table 5. Accuracy experiments with varying numbers of predictions over time. Time was not a factor in the A3 experiment shown in the top row.

ID	acc_at_1	acc_at_2	acc_at_3	acc_at_4	acc_at_5
A3	**0.4744**	**0.6282**	**0.7179**	**0.7905**	**0.8589**
T1	0.4487	0.6239	0.7094	0.7692	0.8076
T2	0.4487	0.6111	0.7008	0.7692	0.8247
T3	0.3846	0.5940	0.7051	0.7564	0.8076

Finally, Table 5 demonstrates that none of the alternatives provided to account for the timestamps (T1, T2, and T3) improved the outcomes of the time trials. These plainly performed worse than an architecture (A3) with a comparable setup but no timestamps. Intuitively, we thought that included the timestamps would aid in understanding the temporal patterns in the activities, however the data demonstrates that this is not the case for the behaviour modelling challenge. We anticipate that chronological information will be significantly more useful in the action credit challenge.

6. Conclusion

Actions, activities, within-activity behaviour, and cross-activity behaviour are all used to explain user conduct in this paper's suggested multi-level conceptual model. We have provided an LSTM-based deep learning architecture that predicts interactive behaviour using this conceptual paradigm. Our design provides a probabilistic model for anticipating user behaviour and spotting out-of-the-ordinary patterns of interaction. We have examined the performance of several designs for varying numbers of action predictions. Our goal for the future is to learn more about how temporal and geographical data might be included into the design of the statistical model.

Our goal is to use convolutional neural networks to simulate the many action n-grams seen in interactive behaviour. Using this method, we may evaluate the relative merits of several deep neural model-based sequence modelling strategies. Starting with the activity detection job, we want to expand our efforts to create structures that address various facets of the suggested multilayer conceptual model. Insights from this paper's suggested deep learning architectures will be used to inform the development of an activity recognition system.

References

[1] Valai Ganesh, S., Agarwal, M., Gupta, S. K., & Rajakarunakaran, S. (2021). Static and dynamic activities prediction of human using machine and deep learning models. In *Innovations in Computer Science and Engineering: Proceedings of 8th ICICSE* (pp. 1-7). Springer Singapore.

[2] Zhang, W., & Li, W. (2019, June). A deep reinforcement learning based human behaviour prediction approach in smart home environments. In *2019 International Conference on Robots & Intelligent System (ICRIS)* (pp. 59-62). IEEE.

[3] Rehder, E., Wirth, F., Lauer, M., & Stiller, C. (2018, May). Pedestrian prediction by planning using deep neural networks. In *2018 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5903-5908). IEEE.

[4] Kamalesh, M. D., & Bharathi, B. (2022). Personality prediction model for social media using machine learning Technique. *Computers and Electrical Engineering*, *100*, 107852.

[5] Ren, H., Song, Y., Wang, J., Hu, Y., & Lei, J. (2018, November). A deep learning approach to the citywide traffic accident risk prediction. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 3346-3351). IEEE.

[6] Wu, J., Abrar, S. M., Awasthi, N., Frias-Martinez, E., & Frias-Martinez, V. (2022). Enhancing short-term crime prediction with human mobility flows and deep learning architectures. *EPJ data science*, *11*(1), 53.

[7] Padmaja, B., Myneni, M. B., & Krishna Rao Patro, E. (2020). A comparison on visual prediction models for MAMO (multi activity-multi object) recognition using deep learning. *Journal of Big Data*, *7*(1), 1-15.

[8] Dang, H. Q., Fürnkranz, J., Biedermann, A., & Hoepfl, M. (2017, October). Time-to-lane-change prediction with deep learning. In *2017 ieee 20th international conference on intelligent transportation systems (itsc)* (pp. 1-7). IEEE.

[9] Nielsen, A. (2019). *Practical time series analysis: Prediction with statistics and machine learning*. O'Reilly Media.

[10] Castro, D., Hickson, S., Bettadapura, V., Thomaz, E., Abowd, G., Christensen, H., & Essa, I. (2015, September). Predicting daily activities from egocentric images using deep learning. In *proceedings of the 2015 ACM International symposium on Wearable Computers* (pp. 75-82).

[11] Liang, T., Zeng, B., Liu, J., Ye, L., & Zou, C. (2018). An unsupervised user behaviour prediction algorithm based on machine learning and neural network for smart home. *IEEE Access*, *6*, 49237-49247.

[12] Opoku Asare, K., Terhorst, Y., Vega, J., Peltonen, E., Lagerspetz, E., & Ferreira, D. (2021). Predicting depression from smartphone behavioural markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. *JMIR mHealth and uHealth*, *9*(7), e26540.

[13] Almeida, A., Bermejo, U., Bilbao, A., Azkune, G., Aguilera, U., Emaldi, M., ... & Arganda-Carreras, I. (2022). A Comparative Analysis of Human Behaviour Prediction Approaches in Intelligent Environments. *Sensors*, *22*(3), 701.

[14] Chopade, P., Khan, S. M., Edwards, D., & von Davier, A. (2018, October). Machine learning for efficient assessment and prediction of human performance in collaborative learning environments. In *2018 IEEE International Symposium on Technologies for Homeland Security (HST)* (pp. 1-6). IEEE.

[15] Bergelin, V. (2017). Human activity recognition and behavioural prediction using wearable sensors and deep learning.