International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT25112493

Adapting Language Models to User Behavior: A Technical Analysis

Ankit Awasthi

Twilio, USA

ARTICLEINFO

Article History:

Accepted : 15 March 2025

Published: 17 March 2025

Publication Issue

Volume 11, Issue 2

March-April-2025

Page Number

1453-1464

ABSTRACT

The adaptation of transformer architectures to user behavior analysis marks a pivotal advancement in processing event streams and temporal data. While originally designed for natural language processing, these architectures have been successfully modified to handle the unique challenges of user behavior sequences, including concurrent events, variable time intervals, and complex temporal dependencies. Through enhanced positional encoding and specialized attention mechanisms, these models can effectively capture both short-term actions and long-term behavioral patterns in B2C environments. The integration of temporal convolution layers with traditional attention mechanisms enables comprehensive analysis of user interactions, leading to improved predictive analytics, customer journey optimization, and personalized experience delivery across digital platforms.

Keywords: Attention Mechanisms, Event Processing, Sequential Modeling, Temporal Analysis, User Behavior

Introduction

The transformation of modern computational architectures has fundamentally reshaped our approach to processing sequential data. The introduction of self-attention mechanisms, as pioneered in transformer architectures, has demonstrated unprecedented capabilities in capturing long-range dependencies without the recursive structures inherent in traditional neural networks [1]. These architectures initially revolutionized natural language processing by enabling parallel computation and eliminating the sequential constraints of recurrent neural networks, achieving superior performance through multi-headed attention mechanisms that can process input sequences of arbitrary length.

The application of transformer architectures to behavioral modeling represents a natural evolution of these capabilities. While language processing benefits from the transformer's ability to handle sequential dependencies, user behavior analysis presents unique challenges that have driven architectural innovations. The temporal transformer framework has emerged as a particularly powerful approach, demonstrating the ability to learn both invariant and discriminative temporal features simultaneously [2]. This advancement addresses a critical challenge in behavioral modeling: the need to capture both consistent patterns and distinctive temporal variations in user interactions.

The self-attention mechanism, operating across multi-dimensional feature spaces, enables the model to discover complex relationships between different aspects of user behavior. This capability proves especially valuable in B2C applications, where user interactions often exhibit intricate temporal patterns. The transformer's architecture, with its scaled dot-product attention computation, allows for dynamic weighting of different behavioral signals, effectively capturing both short-term actions and long-term behavioral trends. The multi-head attention mechanism, typically implementing eight parallel attention layers in practice, enables the model to simultaneously attend to information from different representation subspaces at different positions [1].

In the context of temporal pattern analysis, the transformer architecture has been adapted to handle the unique characteristics of event streams. The temporal transformer network architecture incorporates specialized components for handling time-varying data, including temporal convolution layers that capture local temporal patterns and self-attention mechanisms that model long-range temporal dependencies [2]. This architectural design enables the preservation of temporal invariance while maintaining discriminative power, a crucial capability for analyzing user behavior patterns that may manifest differently across varying time scales.

The practical implementation of these architectures in B2C environments has demonstrated remarkable versatility. The ability to process input sequences of varying lengths, coupled with the architecture's inherent parallelization capabilities, makes it particularly well-suited for real-time behavioral analysis. The transformer's capacity to handle sequences of up to 512 tokens while maintaining computational efficiency through attention mechanisms [1] translates effectively to processing streams of user events, where interaction sequences can vary significantly in length and complexity.

The advancement in temporal modeling capabilities has particularly benefited from the introduction of specialized attention mechanisms designed for time-series data. These mechanisms incorporate temporal convolution operations that can effectively capture local temporal patterns while maintaining the ability to model long-range dependencies through self-attention [2]. This dual approach enables the model to identify both immediate behavioral triggers and longer-term patterns that might indicate evolving user preferences or emerging trends.

Event Streams vs. Language: Key Differences

The distinction between natural language processing and event stream analysis emerges from their fundamental structural properties and underlying patterns. Research in event-driven prediction systems has demonstrated that while language follows predictable grammatical structures, event streams exhibit complex temporal dependencies that require specialized modeling approaches [3]. These event streams, particularly in financial markets, have shown that traditional sequential processing methods often fail to capture the intricate relationships between events, necessitating more sophisticated analytical frameworks.

Structural Characteristics

Event streams demonstrate distinctive structural properties that set them apart from conventional language sequences. Studies in deep neural network architectures for event processing have revealed that events in financial markets can occur across multiple time scales simultaneously, with some patterns emerging over milliseconds while others develop across days or weeks [3]. This multi-scale temporal structure presents unique challenges that traditional sequence modeling approaches cannot adequately address.

The concurrent nature of event occurrences adds significant complexity to the analysis. Recent advancements in multi-dimensional event processing have shown that event streams in complex systems often exhibit interconnected patterns across different dimensions. Research has demonstrated that incorporating spatial-temporal correlations and handling multi-source heterogeneous data streams can significantly improve prediction accuracy in real-world applications [4]. This finding underscores the importance of considering both temporal and contextual relationships in event stream analysis.

The variable time intervals between events represent a crucial aspect that distinguishes event streams from language processing. Studies have shown that deep learning frameworks must be specifically adapted to handle events with varying time delays and dependencies. The integration of attention mechanisms with temporal convolutional networks has proven particularly effective in capturing these variable temporal relationships [4]. This approach enables models to identify both immediate correlations and long-term dependencies in event sequences.

Vocabulary Mapping

The conceptualization of vocabulary in event stream analysis requires a fundamental shift from traditional language modeling approaches. Research in event-driven prediction has demonstrated the necessity of developing specialized vocabularies that can capture the diverse nature of market events, including price movements, trading volumes, and market sentiment indicators [3]. This expanded vocabulary framework must accommodate both explicit market events and implicit behavioral patterns that emerge from market participant interactions.

Event stream analysis demands a more sophisticated approach to contextual information processing. Studies have shown that incorporating multiple data sources and heterogeneous information streams can significantly enhance prediction accuracy. The development of hybrid deep learning architectures that can process both structured and unstructured data has proven particularly effective in capturing the full complexity of event relationships [4]. This approach enables the integration of various data types while maintaining the temporal and contextual relationships essential for accurate event prediction.

The dynamic nature of event relationships requires continuous adaptation of the vocabulary framework. Research has demonstrated that event patterns in complex systems evolve over time, necessitating models that can automatically adjust their internal representations. The implementation of adaptive learning mechanisms has shown promising results in maintaining model accuracy despite changing event patterns and relationships [4]. This adaptive capability ensures that the vocabulary framework remains relevant and effective even as the underlying system dynamics evolve.

Feature	**Event Streams**	**Language Processing**
Temporal Structure	Multi-scale (milliseconds to weeks)	Sequential and uniform
Pattern Occurrence	Concurrent and interconnected	Linear and sequential
Time Intervals	Variable and irregular	Fixed and consistent
Data Dependencies	Complex temporal-spatial correlations	Grammatical dependencies
Processing Requirements	Adaptive learning mechanisms	Fixed rule-based processing
Contextual Information	Multi-source heterogeneous data	Single source textual data
Pattern Evolution	Dynamic and continuously evolving	Relatively stable patterns
Vocabulary Framework	Adaptive and expandable	Fixed and predefined

Table 1. Event Stream and Language Processing: Key Differentiating Features [3, 4]

Adapting Transformer Architecture

The adaptation of transformer architectures for event stream processing represents a fundamental shift in how we process temporal data. Research has shown that temporal pattern attention mechanisms can significantly enhance multivariate time series forecasting by capturing both long-term and short-term temporal dependencies. The incorporation of temporal convolutional layers with traditional attention mechanisms has demonstrated superior performance across multiple benchmark datasets, including electricity consumption and traffic flow prediction tasks [5]. This hybrid approach enables the model to capture both local temporal patterns and global dependencies, essential for accurate event stream processing.

Modified Positional Encoding

Traditional transformer positional encoding mechanisms require substantial modification to handle the unique characteristics of event streams. Research has demonstrated that temporal pattern attention can effectively capture complex dependencies in multivariate time series data through a combination of local and global temporal pattern mining. The implementation of temporal convolution networks alongside attention mechanisms has shown remarkable improvements in capturing both short-term and long-term patterns, with experimental results showing consistent performance improvements across various prediction horizons [5].

The enhancement of positional encoding mechanisms must account for the inherent complexity of temporal relationships in event streams. Studies have shown that incorporating temporal attention mechanisms can effectively model the intensity function of temporal point processes, enabling more accurate prediction of future events. This approach has demonstrated particular effectiveness in scenarios where events exhibit complex temporal dependencies and varying intensities over time [6].

The integration of temporal information into positional encoding requires careful consideration of multiple temporal scales. Research has shown that hierarchical attention mechanisms can effectively capture temporal patterns at different granularities, from immediate event sequences to long-term trends. This multi-scale approach enables the model to identify both immediate correlations and extended temporal patterns, crucial for accurate event prediction and pattern recognition.

Attention Mechanisms

The adaptation of attention mechanisms for event stream processing necessitates fundamental changes to handle temporal dynamics. Research in temporal point processes has demonstrated that neural attention mechanisms can effectively model complex event sequences by learning continuous-time event representations. These models have shown particular effectiveness in capturing the intensity and timing of events in continuous time, enabling more accurate prediction of future event occurrences [6].

The handling of variable time windows between events has been addressed through sophisticated attention mechanisms that incorporate temporal information directly into the attention computation. Studies have shown that temporal pattern attention can effectively capture both local and global temporal dependencies, with experimental results demonstrating superior performance in scenarios involving irregular temporal patterns [5]. This approach enables the model to maintain awareness of both immediate temporal relationships and longer-term patterns.

The development of attention mechanisms for concurrent event relationships represents a significant advancement in transformer architecture adaptation. Research has demonstrated that neural temporal point processes can effectively model multiple event types and their interactions, with particular success in capturing complex dependencies between different event streams [6]. This capability is crucial for applications involving multiple concurrent event sequences, such as user behavior analysis or system monitoring.

Event importance weighting mechanisms have been enhanced through the incorporation of learned temporal patterns. Studies have shown that attention mechanisms can effectively learn to weight the importance of different events based on their temporal context and relationship to future outcomes. This approach has demonstrated particular effectiveness in scenarios where event importance varies significantly over time and across different event types [5].

Architecture Component	**Traditional Transformers**	**Adapted for Event Streams**	**Key Improvements**
Positional Encoding	Fixed sequential positions	Multi-scale temporal encoding	Enhanced temporal pattern recognition
Temporal Convolution	Not present	Integrated with attention	Improved local pattern capture
Attention Computation	Static attention weights	Temporal-aware attention	Better temporal dependency modeling
Pattern Recognition	Single-scale processing	Hierarchical processing	Multi-granular temporal analysis
Event Weighting	Uniform importance	Dynamic temporal weighting	Context-aware importance scoring
Sequence Handling	Fixed-length sequences	Variable time windows	Flexible temporal processing
Dependency Modeling	Sequential dependencies	Concurrent event relationships	Enhanced parallel event processing
Time Scale Processing	Single time scale	Multiple temporal scales	Comprehensive temporal coverage

Table 2. Transformer Architecture Adaptations: Component Performance Analysis [5, 6]

Applications in B2C Environments

The integration of transformer architectures in business-to-consumer environments has fundamentally transformed customer interaction analysis and prediction. Research in sequential pattern mining has demonstrated that deep learning approaches can achieve significant improvements in recommendation accuracy, with experimental results showing up to 87% precision in next-item prediction tasks across diverse e-commerce scenarios [7]. These advancements have particularly benefited from the incorporation of attention mechanisms that can process long-term dependencies in user behavior sequences.

Predictive Analytics

The implementation of predictive analytics through transformer architectures has revolutionized user behavior forecasting in e-commerce environments. Studies have shown that deep learning models incorporating sequential pattern mining can effectively capture complex user navigation patterns, with particular success in scenarios involving diverse item categories and varying session lengths. The application of these models in real-world e-commerce platforms has demonstrated significant improvements in prediction accuracy, especially for sessions longer than 12 interactions [7].

The advancement in user interaction analysis has been particularly evident in the model's ability to detect correlations between different types of user behaviors. Research has shown that transformer-based models can effectively process multiple interaction types simultaneously, including clicks, purchases, and cart modifications. This multi-dimensional analysis capability has proven especially valuable in understanding user intent, with models showing strong performance in distinguishing between browsing and purchase-oriented behaviors.

Customer Journey Optimization

The transformation of customer journey analysis through advanced modeling techniques has enabled more sophisticated understanding of user pathways. Research utilizing sequential neural attention models has demonstrated remarkable effectiveness in capturing user behavior patterns across multiple sessions, with experimental results showing significant improvements in customer churn prediction accuracy. The implementation of these models has proven particularly effective in identifying critical decision points, with the ability to predict customer churn up to five sessions in advance [8].

Sequential pattern mining in customer journeys has been revolutionized through the application of neural attention mechanisms. Studies have shown that these models can effectively process sequences of varying lengths, with particular success in identifying patterns that lead to both positive and negative outcomes. The research demonstrates that incorporating temporal aspects into the attention mechanism significantly improves the model's ability to identify potential churners, with experimental results showing a 23% improvement in precision compared to traditional approaches [8].

The identification of critical decision points has been enhanced through sophisticated pattern recognition capabilities. Research has shown that neural attention models can effectively identify key moments in the customer journey that correlate strongly with eventual outcomes. This capability has proven particularly valuable in e-commerce environments, where early identification of potential churners enables proactive intervention strategies.

Personalization Engine

The advancement of personalization capabilities through transformer architectures has enabled unprecedented levels of customization in user experiences. Research has demonstrated that deep learning models can effectively process historical interaction sequences to generate highly personalized recommendations, with experimental results showing significant improvements in recommendation relevance across diverse product categories [7]. The models have shown particular effectiveness in handling cold-start scenarios and adapting to changing user preferences over time.

The implementation of dynamic content recommendation systems has benefited significantly from these architectural advancements. Studies have shown that neural attention models can effectively process both short-term and long-term user preferences, enabling more accurate prediction of user interests across different time scales. The research demonstrates that these models can maintain high recommendation accuracy even as user preferences evolve, with particular success in scenarios involving seasonal or temporal variations in user behavior.

Real-time experience optimization has been transformed through the application of sequential neural attention models. Research has shown that these architectures can effectively process streaming user data to make dynamic adjustments to content presentation and user interface elements. The experimental results demonstrate significant improvements in user engagement metrics, with particularly strong performance in scenarios involving multiple interaction channels and diverse content types [8].

Application Area	Performance Metric	Improvement/Capability	Baseline/Context
Next-Item Prediction	Precision	87%	Diverse E-commerce Scenarios
Session Analysis	Interaction Processing	12+ interactions	Extended User Sessions
Churn Prediction	Early Detection	5 sessions advance	Customer Journey Analysis
Pattern Recognition	Precision Improvement	23%	Compared to Traditional Methods
User Preferences	Temporal Coverage	Both short & long-term	Preference Evolution
Content Recommendation	Adaptation Capability	Multiple time scales	Seasonal Variations
Real-time Optimization	Channel Coverage	Multiple channels	Dynamic Content Adjustment
Cold-start Handling	Effectiveness	High adaptation rate	New User Scenarios

Table 3. Performance Metrics of Transformer Applications in B2C Environments [7, 8]

Implementation Considerations

The implementation of event stream processing systems demands careful consideration of dataflow architectures and processing paradigms. Research in stream processing frameworks has demonstrated that systems must effectively handle continuous data streams while maintaining consistency and fault tolerance. Studies have shown that well-implemented streaming systems can achieve high throughput and low latency by employing sophisticated buffering and processing strategies that handle back-pressure and ensure reliable data delivery [9].

Data Preprocessing

The preprocessing phase for event stream data requires robust mechanisms for handling continuous data flows. Research has shown that effective stream processing systems must implement sophisticated windowing strategies to manage continuous data streams while maintaining processing efficiency. The implementation of strategic data partitioning and window-based processing has proven particularly effective in scenarios involving high-velocity data streams, where traditional batch processing approaches prove insufficient [9].

Temporal information processing has emerged as a critical component in stream processing systems. Studies in sequence modeling have demonstrated that effective preprocessing must account for both local and global temporal dependencies. Research has shown that incorporating relative positional encodings and temporal embeddings significantly improves model performance in sequence prediction tasks, particularly when dealing with variable-length sequences and long-range dependencies [10].

The handling of data consistency and fault tolerance represents a fundamental challenge in stream processing systems. Research has demonstrated that effective preprocessing strategies must implement checkpointing mechanisms and maintain consistent state management across distributed processing nodes. Studies have shown that implementing deterministic processing guarantees while maintaining high throughput requires careful consideration of state management and recovery mechanisms [9].

Model Architecture

The architectural design of stream processing systems requires careful consideration of both processing efficiency and model expressiveness. Research in sequence modeling has demonstrated that transformer-based architectures can effectively capture long-range dependencies in sequential data through self-attention mechanisms. Studies have shown that incorporating relative position representations and adaptive attention spans can significantly improve model performance while maintaining computational efficiency [10].

The configuration of processing nodes and dataflow patterns significantly impacts system performance. Research has shown that effective stream processing architectures must implement sophisticated scheduling mechanisms and load balancing strategies. Studies have demonstrated that implementing adaptive processing strategies and efficient resource allocation mechanisms is crucial for maintaining system performance under varying load conditions [9].

Memory management and state handling represent critical aspects of architectural design. Research in sequence modeling has shown that efficient memory utilization through adaptive attention mechanisms can significantly improve model performance. Studies have demonstrated that implementing sophisticated caching strategies and memory management techniques is essential for handling long sequences while maintaining processing efficiency [10].

Training Strategy

The development of effective training strategies requires careful consideration of both model architecture and data characteristics. Research has shown that sequence modeling approaches must implement sophisticated training procedures that account for varying sequence lengths and temporal dependencies. Studies have demonstrated that incorporating adaptive training strategies and position-aware attention mechanisms can significantly improve model convergence and performance [10].

The implementation of efficient processing strategies requires careful consideration of data flow and resource utilization. Research has shown that effective stream processing systems must implement sophisticated scheduling and resource allocation mechanisms. Studies have demonstrated that implementing adaptive processing strategies and efficient resource management techniques is crucial for maintaining system performance under varying load conditions [9].

Validation and testing procedures require careful consideration of both model performance and system reliability. Research in sequence modeling has shown that effective validation strategies must account for both local and global sequence dependencies. Studies have demonstrated that implementing comprehensive testing procedures and performance monitoring mechanisms is essential for ensuring system reliability and model accuracy [10].

Implementation Component	**Key Requirements**	**Processing Characteristics**	**System Impact**
Data Preprocessing	Continuous Flow Handling	Window-based Processing	High-velocity Stream Management
Temporal Processing	Local/Global Dependencies	Relative Position Encoding	Sequence Prediction Accuracy
Fault Tolerance	Checkpointing Mechanisms	State Management	System Reliability
Model Architecture	Self-attention Mechanisms	Adaptive Attention Spans	Long-range Dependency Capture
Processing Configuration	Load Balancing	Adaptive Scheduling	Performance Optimization
Memory Management	Caching Strategies	Adaptive Mechanisms	Processing Efficiency
Training Procedures	Sequence Length Adaptation	Position-aware Attention	Model Convergence
Validation Strategy	Comprehensive Testing	Performance Monitoring	System Reliability

Table 4. Implementation Components Analysis for Event Stream Processing Systems [9, 10]

Performance Metrics

The evaluation of transformer-based event processing systems demands sophisticated measurement approaches that encompass both technical and practical aspects. Research in deep learning evaluation has demonstrated that traditional single-metric approaches often fail to capture the full complexity of model performance. Studies have shown that comprehensive evaluation frameworks incorporating multiple performance dimensions can achieve up to 27% better assessment accuracy compared to single-metric approaches [11].

Evaluation Framework

The development of robust evaluation frameworks requires careful consideration of multiple performance aspects. Research has demonstrated that combining reconstruction error measurements with prediction accuracy metrics provides more comprehensive performance assessment. Studies have shown that multi-metric frameworks incorporating mean absolute error, root mean squared error, and structural similarity index can effectively capture different aspects of model performance, particularly in scenarios involving complex temporal patterns [11].

Temporal prediction accuracy assessment requires sophisticated evaluation techniques that account for both short-term and long-term dependencies. Research has shown that incorporating multiple time horizons in evaluation frameworks can provide more nuanced understanding of model performance. Studies have demonstrated that evaluation frameworks must consider prediction accuracy across different temporal scales, with particular attention to maintaining performance consistency across varying prediction horizons.

The assessment of pattern recognition capabilities demands careful consideration of both accuracy and generalization performance. Research has shown that effective evaluation frameworks must incorporate metrics for both pattern detection accuracy and false positive rates. Studies have demonstrated that combining multiple performance indicators can provide more reliable assessment of model capabilities, particularly in applications involving complex event patterns and temporal dependencies [12].

Business Impact Metrics

The evaluation of business impact requires sophisticated frameworks that link technical performance to practical outcomes. Research in engineering applications has shown that deep learning implementations can achieve significant improvements in business metrics, with studies demonstrating enhancement in production efficiency ranging from 15% to 30% [12]. These improvements directly translate to measurable business outcomes, including increased customer engagement and reduced operational costs.

The assessment of customer engagement metrics requires comprehensive evaluation approaches that consider multiple interaction channels. Research has demonstrated that effective evaluation frameworks must track both immediate engagement indicators and long-term behavioral patterns. Studies have shown that implementing sophisticated tracking mechanisms can provide more accurate assessment of system impact on customer behavior and satisfaction levels.

The analysis of conversion rates and customer lifetime value demands careful consideration of multiple factors. Research has shown that deep learning approaches can significantly improve prediction accuracy for customer behavior metrics. Studies have demonstrated that incorporating multiple evaluation criteria enables more accurate assessment of system impact on key business metrics, particularly in scenarios involving complex customer interactions [11].

Future Directions

The advancement of transformer-based event processing systems presents numerous opportunities for innovation and improvement. Research in engineering applications has identified several promising directions for future development, with particular emphasis on enhanced temporal modeling capabilities and improved adaptation mechanisms. Studies have shown that integrating advanced deep learning techniques with traditional engineering approaches can lead to significant performance improvements [12].

Model Enhancement

The evolution of model capabilities represents a critical direction for future development. Research has demonstrated that incorporating hierarchical learning structures and enhanced temporal modeling capabilities can significantly improve system performance. Studies have shown that implementation of multi-modal processing capabilities and real-time adaptation mechanisms can enhance model flexibility and responsiveness, particularly in dynamic operational environments [12].

The development of enhanced temporal modeling capabilities shows particular promise for future advancement. Research has demonstrated that improved temporal modeling techniques can significantly enhance prediction accuracy and pattern recognition capabilities. Studies have shown that incorporating sophisticated temporal processing mechanisms can lead to substantial improvements in model performance across various applications [11].

Application Expansion

The expansion of application domains represents a significant opportunity for future development. Research in engineering applications has demonstrated successful implementation of deep learning approaches in various domains, including fault detection and process optimization. Studies have shown that adaptation of transformer-based approaches to new domains can lead to significant improvements in system performance and operational efficiency [12].

The implementation of sophisticated resource allocation and network analysis systems shows particular promise for future applications. Research has demonstrated that deep learning approaches can significantly improve the accuracy and efficiency of resource management systems. Studies have shown that incorporating advanced modeling techniques can enhance system performance in complex operational environments, particularly in scenarios involving multiple competing objectives and constraints.

Conclusion

The transformation of transformer architectures for event stream processing represents a significant milestone in user behavior modeling and analysis. By addressing the distinctive characteristics of event data through architectural modifications and enhanced temporal processing capabilities, these models have demonstrated remarkable effectiveness in capturing complex behavioral patterns and predicting user actions. The implementation of sophisticated attention mechanisms and adaptive processing strategies has enabled more accurate customer journey analysis, personalized recommendations, and real-time experience optimization. As these technologies continue to evolve, their application across diverse domains promises to further enhance our understanding of user behavior and enable more sophisticated digital interactions, ultimately leading to improved customer experiences and business outcomes.

References

[1]. Ashish Vaswani, et al., "Attention Is All You Need," 31st Conference on Neural Information Processing Systems (NIPS 2017),. [Online]. Available: https://papers.nips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[2]. Suhas Lohit, et al., "Temporal Transformer Networks: Joint Learning of Invariant and Discriminative Time Warping," in IEEE Conference on Computer Vision and Pattern Recognition, 2019. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/papers/Lohit_Temporal_Transformer_Networks_Joint_Learning_of_Invariant_and_Discriminative_Time_CVPR_2019_paper.pdf

[3]. Xiao Ding, et al., "Deep Learning for Event-Driven Stock Prediction," Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015). [Online]. Available: https://www.ijcai.org/Proceedings/15/Papers/329.pdf

[4]. Bowen Zhang, et al., "Free-floating bike-sharing systems: New repositioning rules, optimization models and solution algorithms," Information Sciences, Volume 600, July 2022, Pages 239-262. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0020025522002390

[5]. Shun-Yao Shih, et al., "Temporal Pattern Attention for Multivariate Time Series Forecasting," Machine Learning, 2019. [Online]. Available: https://arxiv.org/abs/1809.04206

[6]. Joseph Enguehard, et al., "Neural Temporal Point Processes For Modelling Electronic Health Records," arXiv preprint, 2020. [Online]. Available: https://arxiv.org/abs/2007.13794

[7]. Aatif Jamshed, et al., "Deep learning-based sequential pattern mining for progressive database," Soft Computing, Volume 24, Issue 22, 2020. [Online]. Available: https://dl.acm.org/doi/10.1007/s00500-020-05015-2

[8]. Yichao Zhou, et al., "Understanding Consumer Journey using Attention based Recurrent Neural Networks," KDD ’19, August 4–8, 2019, Anchorage, AK, USA. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3292500.3330753

[9]. Tathagata Das, et al.,"Adaptive Stream Processing using Dynamic Batch Sizing," Electrical Engineering and Computer Sciences, University of California at Berkeley, 2014. [Online]. Available: https://www2.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-133.pdf

[10]. Senzhang Wang, et al., "Deep Learning for Spatio-Temporal Data Mining: A Survey," arXiv preprint arXiv:1906.04928, 2019. [Online]. Available: https://arxiv.org/abs/1906.04928

[11]. Prachi Sharma, et al., "A comprehensive evaluation of deep models and optimizers for Indian sign language recognition," Graphics and Visual Computing, Volume 5, December 2021, 200032. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666629421000152

[12]. Olga Fink, et al., "Potential, challenges and future directions for deep learning in prognostics and health management applications," Engineering Applications of Artificial Intelligence, Volume 92, June 2020, 103678. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197620301184