Deep Learning for Multivariate Time Series Imputation: A Survey

www.surveyx.cn

Abstract

This survey paper provides a comprehensive examination of the transformative role of deep learning in the imputation of multivariate time series data, focusing on its application across various domains such as healthcare, finance, and industrial processes. The paper underscores the critical importance of handling missing data to ensure the accuracy and reliability of predictive models. It explores advanced neural network architectures, including Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Transformer models, and hybrid architectures, each evaluated for their efficacy in reconstructing incomplete datasets. The survey highlights the superiority of deep learning approaches over traditional imputation techniques, particularly in capturing complex temporal dependencies and enhancing predictive accuracy. It also discusses the challenges of scalability, computational efficiency, and model interpretability, emphasizing the need for ongoing research to develop more adaptive and robust models. Case studies across various sectors illustrate the practical implications and successes of deep learning in improving data reconstruction and analysis. The paper concludes by reaffirming the potential of deep learning to revolutionize multivariate time series imputation, paving the way for more accurate and reliable data-driven insights in diverse applications.

1 Introduction

1.1 Importance of Handling Missing Data

Handling missing data is crucial for the accuracy and reliability of time series analysis and prediction across various domains, including healthcare, finance, and industrial processes. In healthcare, missing data hampers the extraction of meaningful insights, often due to redundant or ambiguous labels [1]. In finance, where volatility forecasting is paramount, incomplete datasets can disrupt trading strategies, necessitating robust imputation methods. The effectiveness of machine learning algorithms is heavily reliant on complete datasets, as missing values can undermine model accuracy and reliability. In industrial contexts, the challenges of data sovereignty and expert knowledge are exacerbated by missing data, underscoring the need for effective handling techniques [2].

Moreover, the limitations of untrained neural networks in reconstructing complex datasets, such as MRI images, highlight the demand for improved methodologies [3]. In business analytics, deep learning models have shown the ability to detect intricate patterns and enhance predictive accuracy, contingent upon the availability of comprehensive data [4]. The rapid evolution of deep learning introduces additional complexities that researchers must navigate to maintain model reliability in real-world applications. Traditional training methods, like backpropagation, can hinder effective data reconstruction, emphasizing the importance of addressing missing data [5].

In educational settings, modeling student interactions for knowledge tracing is directly impacted by missing data, limiting insights into learning processes [6]. Similarly, in brain-computer interface (BCI) research, dataset size limitations stress the necessity for effective data handling techniques

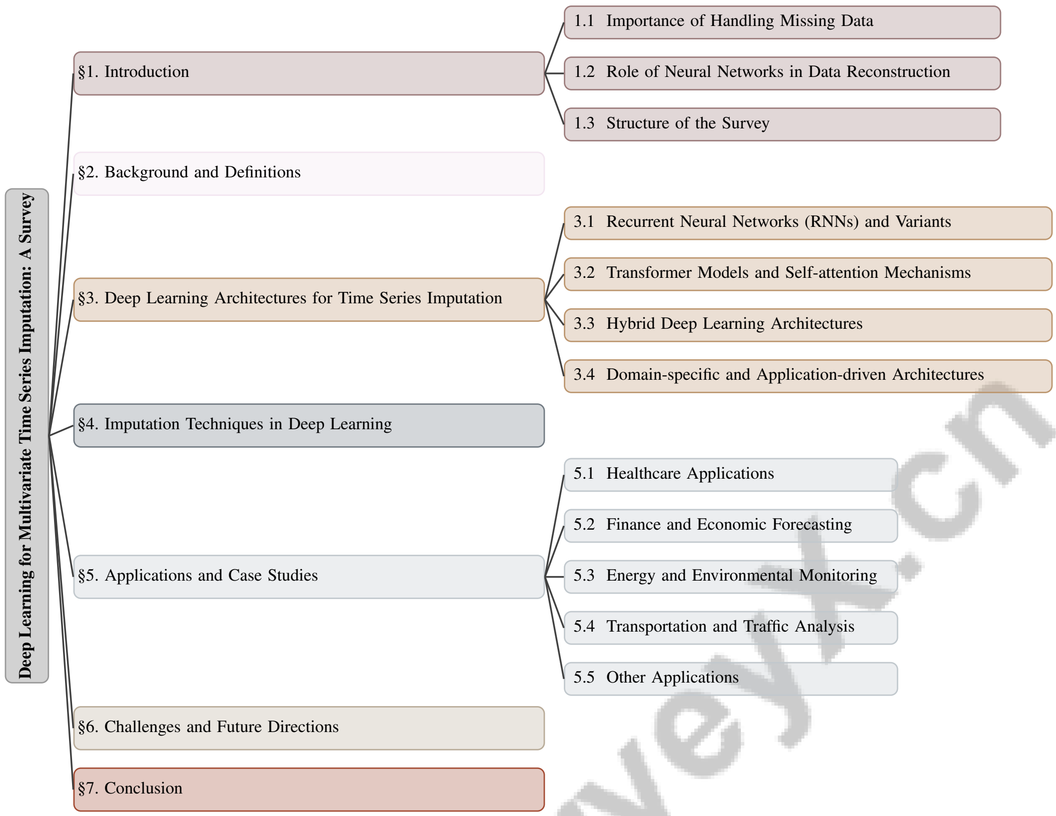


Figure 1: chapter structure

[7]. Neural networks, with their sophisticated architectures, present promising solutions for data reconstruction, yet their success is inherently linked to the quality of input data [8]. Thus, robust handling of missing data is essential across diverse fields to uphold the integrity of time series analyses and enhance the performance of predictive models.

1.2 Role of Neural Networks in Data Reconstruction

Neural networks have become fundamental in reconstructing missing data within multivariate time series, leveraging advanced architectures that enhance data integrity and predictive capabilities. For instance, the ConvlstmCsiNet architecture exemplifies how deep learning can improve feedback compression quality and efficiency in systems like massive MIMO [9]. The TSMixer architecture further captures both temporal and cross-variate information, advancing the reconstruction process [10].

Probabilistic deep learning approaches increase the reliability of neural networks by incorporating uncertainty into models, significantly aiding the reconstruction of missing data [11]. This is particularly relevant in knowledge tracing applications, where self-attention models enhance prediction accuracy by incorporating contextual information and modeling student forget behavior [6]. Despite these advancements, the interpretability of neural networks remains a challenge, often leading to their characterization as 'black boxes' [8].

Recurrent networks, due to their sequential nature, are especially well-suited for time series data reconstruction [12]. Hybrid models, combining CNNs and RNNs, address the complexities of missing data, showcasing the versatility of deep learning architectures [13]. The K-UNN method, which utilizes physical priors, illustrates how neural networks can enhance reconstruction accuracy, further emphasizing their critical role in data reconstruction [3].

In business analytics, deep learning models, including the proposed deep-embedded architecture, consistently surpass traditional models, highlighting their effectiveness in data reconstruction [4]. Neural networks empower more informed decision-making tailored to specific tasks, significantly improving data reconstruction outcomes [14]. Collectively, these advancements underline the essential role of neural networks in addressing the complexities of missing data, facilitating accurate and reliable time series analyses across various domains.

1.3 Structure of the Survey

This survey is systematically structured to provide an in-depth exploration of deep learning approaches for multivariate time series imputation. The introduction highlights the significance of addressing missing data and the pivotal role of neural networks in data reconstruction. Following this, a background section elucidates foundational concepts such as deep learning, multivariate time series, imputation, and the challenges associated with missing data. The subsequent section examines various deep learning architectures utilized in time series imputation, including Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), Transformer models, and hybrid architectures, analyzing their specific advantages and limitations.

The survey then transitions to a comprehensive examination of imputation techniques within deep learning models, ranging from traditional methods like zero and mean imputation to advanced techniques utilizing variational autoencoders. The exploration of application-specific imputation strategies highlights the versatility of these methods across diverse fields, exemplified by innovative approaches such as the transformer-based model NAIM, which effectively addresses missing values in tabular datasets without conventional imputation, and deep learning algorithms that enhance data harmonization by jointly matching, imputing, and transforming features from disparate databases, thereby improving predictive performance and data integration in complex domains like healthcare [15, 1]. Real-world applications and case studies illustrate the practical implications and successes of deep learning in multivariate time series imputation, covering sectors such as healthcare, finance, energy monitoring, and transportation.

The survey thoroughly examines prevailing challenges in the deep learning field, including scalability, interpretability, and computational efficiency, which hinder effective deployment of advanced models. It addresses the limitations of current methodologies, such as the black-box nature of many algorithms and the insufficient attention to structured data modeling. Additionally, targeted future research directions are proposed to tackle these challenges, thereby promoting advancements in deep learning applications across diverse domains [16, 17]. The conclusion synthesizes key findings, reaffirming the transformative impact of deep learning on multivariate time series imputation and its potential for future advancements. This structured approach ensures a holistic understanding of the topic, facilitating a nuanced appreciation of both theoretical and practical dimensions of deep learning in time series data analysis.The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Multivariate Time Series and Imputation

Multivariate time series (MTS) data consist of multiple interrelated time-dependent variables, each reflecting distinct system behaviors over time. This complexity is evident in financial markets, where analyzing multiple currency pairs is crucial for accurate volatility predictions [18]. The intricate relationships within MTS data present challenges for machine and deep learning applications, which must navigate their high-dimensional nature [12]. In industrial settings, maintaining data integrity through effective imputation is essential for optimizing processes and ensuring data sovereignty [2].

Imputation in MTS aims to estimate missing values, enhancing dataset completeness and reliability, which is vital for accurate analysis and forecasting. Missing values can severely hinder data utility, necessitating robust imputation methods. Deep learning, with its proficiency in high-dimensional function estimation and pattern recognition, is a powerful tool for MTS imputation, enabling high-fidelity reconstruction of missing data [12]. Techniques like DoppelGANger, a GAN tailored for generating high-fidelity synthetic time series data, illustrate advancements in imputation for economic indicators such as Treasury yields [19].

The application of deep learning models in MTS imputation is underscored by their ability to capture complex temporal dependencies, as evidenced in benchmarks involving datasets with intricate patterns and interactions [12]. These capabilities are crucial for enhancing predictive accuracy across diverse applications, from financial forecasting to industrial process optimization. Imputation remains pivotal

AI-generated, for reference only.

in effectively utilizing MTS data, facilitating informed decision-making and improved forecasting outcomes.

2.2 Challenges of Missing Data in Time Series

Missing data in time series analysis presents multifaceted challenges, adversely affecting predictive model accuracy and reliability across domains. A primary issue is the reliance on univariate models that overlook cross-variate information, leading to suboptimal performance, particularly in complex systems like financial markets [18]. This challenge is compounded by the complexity of advanced models, which can be difficult for practitioners to implement effectively in business analytics [4]. The interpretability of neural networks poses a significant hurdle, limiting insights into their predictive mechanisms and hindering their utility in imputation tasks [8].

Infrequent events, such as economic recessions, complicate model training due to insufficient data, making accurate predictions difficult [19]. Existing methods for reducing the complexity of high-dimensional dynamical systems often require intrusive access to system operators during the projection step [20]. The absence of network architectures incorporating physical priors negatively impacts reconstruction quality, paralleling the challenges posed by missing data [3].

The inherent locking problem in both biological and artificial neural networks exacerbates these challenges, as feedback cannot be utilized until all computations are completed, hindering efficient data reconstruction [21]. Furthermore, researchers face confusion due to the vast array of options for loss functions and metrics, which can lead to misapplication in various contexts [14]. Existing methods for forecasting volatility often rely on proxies like ranges instead of directly modeling volatility, resulting in inaccuracies and emphasizing the difficulties associated with missing data [18].

Additionally, the computational complexity of probabilistic inference and challenges in accurately modeling uncertainty in data and model parameters present significant obstacles in managing missing data [11]. The limitations of traditional methods, such as structured support vector machines (SVM) and conditional random fields (CRF), which operate under linear assumptions, further constrain their ability to analyze complex, nonlinear relationships present in various data types, including speech signals [22]. Finally, stakeholder reluctance to share sensitive data and the necessity for expert knowledge in designing tailored solutions pose critical barriers to effective data management [2]. These challenges highlight the need for advanced modeling approaches and robust imputation techniques to effectively handle missing data, ensuring predictive models maintain the accuracy and reliability essential for practical applications across diverse fields.

3 Deep Learning Architectures for Time Series Imputation

Deep learning architectures have significantly advanced the field of time series imputation by addressing its inherent complexities. This section delves into the efficacy of Recurrent Neural Networks (RNNs) and their variants, which are pivotal in modeling sequential data and capturing temporal dependencies. Figure 2 illustrates the hierarchical categorization of these deep learning architectures for time series imputation, detailing the primary categories of RNNs and their variants, Transformer Models and Self-attention Mechanisms, Hybrid Deep Learning Architectures, and Domain-specific and Application-driven Architectures. Each category is further broken down into subcategories and specific techniques or models, highlighting their contributions to handling sequential data, capturing temporal dependencies, and improving imputation accuracy across various domains. RNNs, particularly Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), excel in reconstructing missing data, thereby improving the reliability of time series analyses. The following subsections will detail the mechanisms and advantages of these architectures in time series imputation.

3.1 Recurrent Neural Networks (RNNs) and Variants

RNNs, including LSTMs and GRUs, are crucial for time series imputation due to their proficiency in handling sequential data and capturing temporal dependencies, particularly in multivariate contexts where sequence information is vital [18]. These networks maintain sequence information, enhancing data reliability and predictive accuracy.

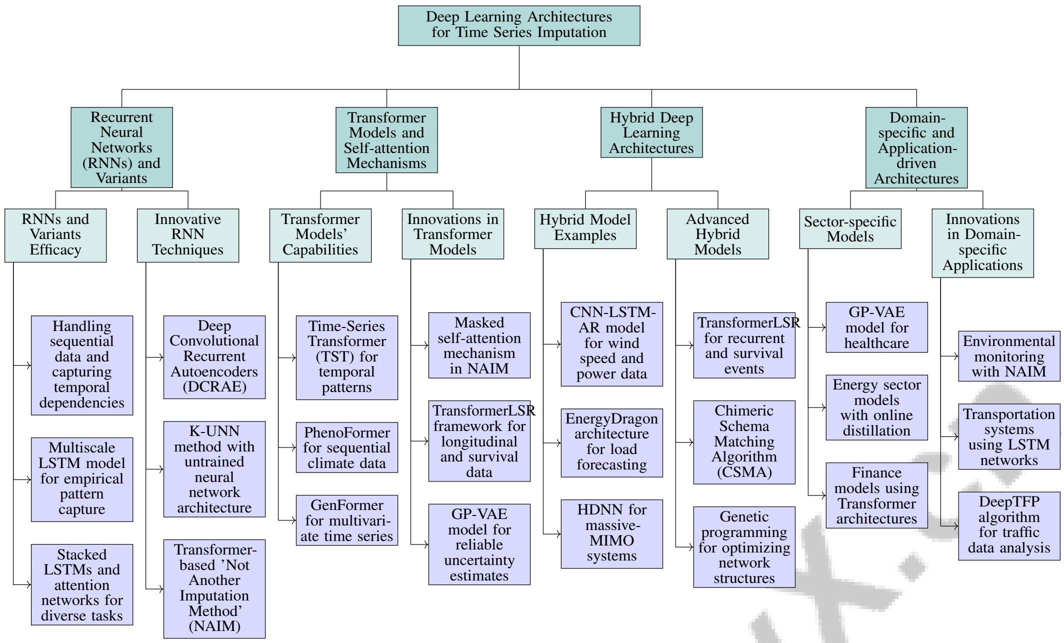


Figure 2: This figure illustrates the hierarchical categorization of deep learning architectures for time series imputation, detailing the primary categories of Recurrent Neural Networks (RNNs) and their variants, Transformer Models and Self-attention Mechanisms, Hybrid Deep Learning Architectures, and Domain-specific and Application-driven Architectures. Each category is further broken down into subcategories and specific techniques or models, highlighting their contributions to handling sequential data, capturing temporal dependencies, and improving imputation accuracy across various domains.

The Deep Convolutional Recurrent Autoencoders (DCRAE) method exemplifies the integration of RNNs with convolutional architectures, using a deep convolutional autoencoder for low-dimensional representations and a modified LSTM for temporal modeling [20]. This synergy showcases RNNs' ability to manage high-dimensional data complexities, facilitating accurate imputation.

Techniques like the multiscale LSTM model effectively capture empirical patterns in Forex data, highlighting the versatility of RNN variants in time series imputation [18]. The exploration of stacked LSTMs and attention networks further emphasizes RNNs' adaptability in diverse imputation tasks [23]. These models are particularly valuable in real-time data processing applications, where efficient imputation methods are critical.

The K-UNN method, using an untrained neural network architecture with physical priors, provides a comparative framework for understanding RNNs' advantages in time series imputation [3]. This survey categorizes research on loss functions and metrics, offering a structured understanding of RNN-based models' strengths and weaknesses [14].

The evolution of RNNs and their variants in time series imputation underscores their pivotal role in advancing data reconstruction methodologies. These architectures employ innovative techniques that surpass traditional methods, enhancing sequential data analysis. For instance, the transformer-based "Not Another Imputation Method" (NAIM) uses feature-specific embeddings and a masked self-attention mechanism to learn from available data without traditional imputation, demonstrating superior predictive performance across multiple datasets. Additionally, a novel deep learning algorithm harmonizes data from diverse sources, effectively matching shared features and imputing unshared variables, proving its effectiveness in complex environments like healthcare. Collectively, these approaches pave the way for sophisticated and robust imputation techniques across various domains, ultimately improving data-driven models' reliability [15, 1].

As illustrated in Figure 3, the categorization of Recurrent Neural Networks (RNNs) and their variants highlights key models, applications, and comparative methods in time series imputation, further emphasizing the diverse landscape of techniques available for enhancing predictive accuracy and data integrity.

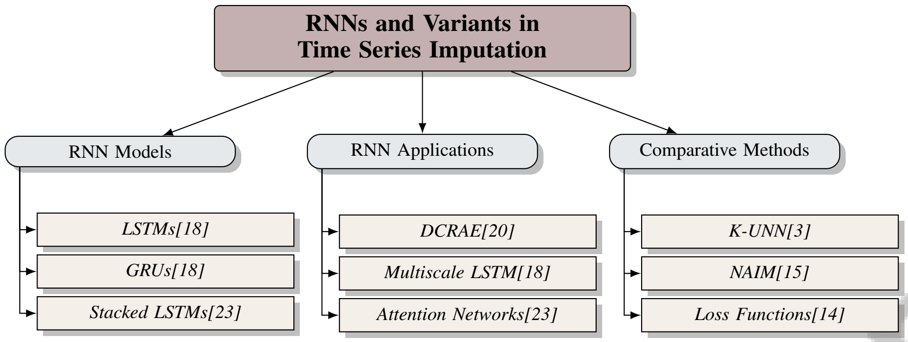


Figure 3: This figure illustrates the categorization of Recurrent Neural Networks (RNNs) and their variants, highlighting key models, applications, and comparative methods in time series imputation.

3.2 Transformer Models and Self-attention Mechanisms

Transformer models and self-attention mechanisms have revolutionized time series imputation by offering innovative approaches to handle complex dependencies and missing data. The Time-Series Transformer (TST) exemplifies the application of transformer architectures in processing segmented patches of flow-volume curves, effectively capturing intricate temporal patterns for imputation [24]. The self-attention mechanism, central to Transformer models, has been leveraged in PhenoFormer to process sequential climate data, enabling feature learning directly from raw inputs [25].

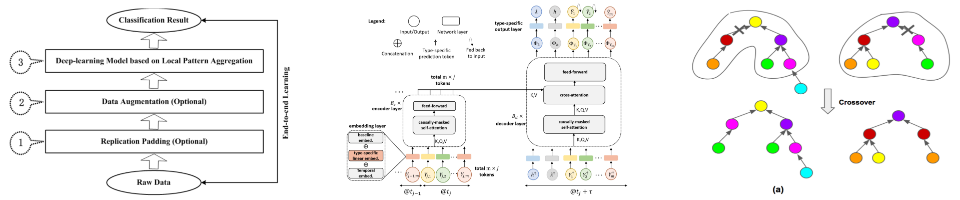
Transformers' ability to manage long-range dependencies is showcased in managing radial spokes, demonstrating their capacity to address complex imputation scenarios [26]. This capability benefits large datasets and long-term dependencies, as evidenced by the GenFormer model, which uses a Transformer-based approach to generate multivariate time series [27]. Such models enhance imputation processes' accuracy and reliability by effectively modeling intricate relationships in multivariate time series data.

The NAIM model introduces a masked self-attention mechanism and feature-specific embeddings, allowing it to learn from incomplete data without traditional imputation methods [15]. This approach improves the model's ability to reconstruct missing data and enhances adaptability across various domains. The introduction of weighted attention networks emphasizes the importance of visualizing attention vectors, enhancing model prediction interpretability [23]. This focus on interpretability is crucial for understanding how Transformer models achieve their predictions, increasing their utility in practical applications.

Transformer models and self-attention mechanisms signify a transformative advancement in time series imputation techniques, effectively managing missing data and complex interdependencies. For instance, the NAIM model employs a transformer-based approach that eliminates the need for conventional imputation, demonstrating superior performance across multiple datasets compared to traditional models. Similarly, the TransformerLSR framework integrates deep temporal point processes to model longitudinal measurements, recurrent events, and survival data, addressing dependencies often overlooked in existing methods. Furthermore, the GP-VAE model employs a deep sequential latent variable approach for multivariate time series imputation, enhancing interpretability and providing reliable uncertainty estimates while outperforming classical methods. Collectively, these innovations illustrate the potential of transformer-based models to enhance predictive accuracy and resilience in the presence of incomplete data [15, 28, 29]. Their application across diverse domains underscores their potential to improve the accuracy and reliability of time series analyses, paving the way for sophisticated and interpretable imputation techniques.

As shown in Figure 4, the example titled "Deep Learning Architectures for Time Series Imputation; Transformer Models and Self-attention Mechanisms" explores advanced methodologies in deep learning, specifically focusing on the application of transformer models and self-attention mechanisms for time series imputation. Illustrated through a series of figures, the example delves into various architectures and techniques pivotal to this field. One image highlights the use of a deep learning

AI-generated, for reference only.



(a) Deep Learning Model for Image Classification[30]

(b) Type-Specific Neural Network for Time Series Prediction[28]

(c) Crossover in Genetic Algorithms[31]

Figure 4: Examples of Transformer Models and Self-attention Mechanisms

model for image classification, showcasing a structured flowchart that outlines the stages from local pattern aggregation to optional data augmentation and replication padding, culminating in classification. Another image presents a type-specific neural network architecture tailored for time series prediction, emphasizing the critical role of input and output layers alongside network layers in processing and predicting type-specific values. Additionally, the crossover operation in genetic algorithms is depicted, illustrating how parent solutions are combined through a tree structure to generate offspring with new genetic combinations, thereby enhancing the evolutionary process. Collectively, these images and descriptions underscore the versatility and sophistication of transformer models and self-attention mechanisms in addressing complex tasks such as time series imputation [30, 28, 31].

3.3 Hybrid Deep Learning Architectures

Hybrid deep learning architectures have emerged as a potent strategy for time series imputation, effectively integrating diverse neural network models' strengths to enhance predictive accuracy and robustness. This approach leverages algorithms, including convolutional and recurrent neural networks, to address the complexities of time-dependent data, surpassing traditional machine learning techniques. As deep learning progresses, its application in time series analysis transforms how practitioners handle large datasets, enabling more reliable predictions across domains like finance, healthcare, and environmental monitoring [12, 13, 17]. These architectures adeptly combine models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks to tackle multivariate time series data challenges.

A notable example of hybrid architecture is the CNN-LSTM-AR model, which integrates CNNs for feature extraction and LSTMs for sequence prediction, effectively mitigating noise challenges in wind speed and power data. This model highlights hybrid architectures' potential to enhance imputation accuracy by capturing spatial and temporal dependencies inherent in data. The EnergyDragon architecture exemplifies leveraging hybrid models in the energy sector by integrating automated deep learning techniques with deep neural networks (DNNs) to enhance load forecasting accuracy. Built upon the DRAGON package, this innovative framework automatically selects relevant features and optimizes network architecture and hyperparameters, addressing the complexities of forecasting electricity consumption influenced by various external factors [32, 33].

In massive-MIMO systems, the hybrid deep neural network (HDNN) architecture integrates analog deep neural networks (ADNN) with digital processing techniques, particularly significant for millimeter-wave (mmWave) communications. This integration allows for efficient design and implementation of hybrid analog-digital signal processing (HSP) systems, enabling the HDNN architecture to approximate complex transmitter and receiver mappings with high precision, facilitating advanced beamforming strategies that achieve performance comparable to fully-digital systems while utilizing fewer radio frequency (RF) chains. This hybrid approach bears conceptual similarities to time series imputation tasks, as both involve effective synthesis and processing of data to enhance system performance [34, 35]. Additionally, the recursive autoencoder model processes sequences to generate lower-dimensional representations, facilitating effective time series imputation.

Advanced hybrid models such as the TransformerLSR stand out by modeling recurrent and survival events as competing temporal point processes. This model employs an innovative trajectory representation and integrates known latent structures among longitudinal variables, effectively addressing the complexities of joint modeling in various fields such as biomedical studies and social sciences. By leveraging hybrid architectures, it captures intricate temporal dynamics and significantly enhances imputation outcomes, as demonstrated by its ability to simultaneously model longitudinal measurements, recurrent events, and survival data while considering their interdependencies. This approach surpasses traditional methods, providing improved predictive performance and interpretability in scenarios with missing data [36, 15, 28, 29].

The Chimeric Schema Matching Algorithm (CSMA) introduces a deep learning-based approach using autoencoders to compress data into a shared latent space, facilitating the matching of features across databases. This method underscores the utility of hybrid architectures in leveraging shared latent spaces for improved imputation [1]. Additionally, the innovation of using genetic programming to evolve recurrent node structures exemplifies the potential of hybrid models to discover and optimize complex network structures for enhanced data imputation [31].

Hybrid deep learning architectures represent a transformative advancement in time series imputation, effectively leveraging the complementary strengths of various models to address the complexities of temporal data. This approach capitalizes on deep learning's ability to automatically recognize intricate patterns and capture temporal dependencies, which are often challenging for traditional techniques reliant on hand-crafted features. By integrating multiple model types, hybrid architectures enhance the accuracy and robustness of imputation techniques while streamlining feature extraction and learning processes, optimizing performance in diverse real-world applications such as financial forecasting and medical data analysis [12, 37]. Their application across various domains highlights their potential to improve the accuracy and reliability of imputation processes, paving the way for more sophisticated and adaptable techniques in time series analysis.

3.4 Domain-specific and Application-driven Architectures

Domain-specific and application-driven architectures in deep learning for time series imputation are meticulously crafted to confront the distinct challenges and requirements inherent to specific fields, such as healthcare and finance, thereby significantly enhancing the accuracy and relevance of imputation methods. Recent advancements like the GP-VAE model demonstrate superior performance in imputing high-dimensional multivariate time series data by leveraging a Gaussian process for smooth temporal evolution and providing interpretable uncertainty estimates. Innovative approaches like the NAIM transformer model eliminate the need for traditional imputation techniques by utilizing feature-specific embeddings and masked self-attention, showcasing improved predictive capabilities across various datasets. These specialized architectures address missing data complexities and optimize the learning process for practical applications, making them invaluable tools in their respective domains [12, 15, 29].

In the energy sector, deep learning models are specifically adapted to forecast energy consumption and production, integrating methodologies like online distillation and residual error learning to refine predictions [33]. These approaches enable models to continuously learn and adapt to new data, improving accuracy in dynamic environments where energy patterns fluctuate due to various factors such as weather and human activity.

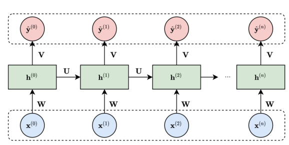
Healthcare applications benefit from architectures that incorporate domain-specific features, such as physiological signals or patient history, to enhance missing data imputation in medical records. These advanced models frequently utilize hybrid architectures that integrate Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), effectively capturing both spatial and temporal dependencies essential for improving diagnosis and treatment planning in complex medical scenarios [12, 30].

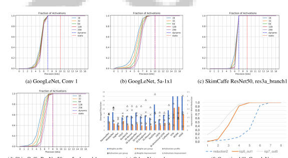
In finance, domain-driven architectures are designed to handle the complexities of economic indicators and stock market data. These models frequently employ Transformer architectures to effectively address long-range dependencies and inherent volatility in financial time series data. By leveraging feature-specific embeddings and a masked self-attention mechanism, these models enhance robustness and reliability in imputation tasks, allowing them to learn from available data without relying on traditional imputation techniques. This approach improves predictive performance and boosts resilience in the presence of missing data, marking a significant advancement in financial analytics [12, 15].

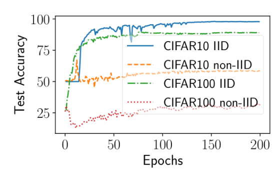
Environmental monitoring applications utilize advanced architectures that integrate diverse sensor data with domain-specific features, facilitating precise imputation of missing environmental parameters. This integration is enhanced by employing techniques such as deep learning, which effectively models complex relationships within the data, even in sparsity. Novel methods like the "Not Another Imputation Method" (NAIM) leverage transformer-based models to address missing values without traditional imputation, improving predictive performance. Frameworks combining zero-bias deep learning with quick event detection algorithms offer reliable solutions for identifying anomalies in environmental data, ensuring accuracy and efficiency in monitoring efforts [38, 15, 39, 40]. These models are essential for tracking climate patterns and assessing ecological changes, where missing data can significantly impact analysis and decision-making.

Transportation systems benefit from application-specific architectures that address traffic flow and vehicle dynamics challenges. Models designed for traffic data analysis frequently utilize Long Short-Term Memory (LSTM) networks to capture the sequential dynamics inherent in traffic patterns, significantly enhancing the accuracy of traffic flow predictions and overall transportation network management. Recent advancements, such as the DeepTFP algorithm, demonstrate that integrating multiple deep learning architectures, including LSTM and residual neural networks, can further improve prediction performance by accounting for various temporal properties and complex influencing factors, optimizing transportation planning and congestion mitigation strategies [41, 42, 31, 23, 43].

Domain-specific and application-driven architectures in deep learning for time series imputation represent a significant advancement, allowing for the development of tailored models that address the unique challenges and requirements of various fields. These architectures leverage advanced methodologies, like variational autoencoders and Gaussian processes, to enhance the accuracy and interpretability of imputation results, ultimately surpassing traditional methods in handling high-dimensional data with missing values. This evolution underscores the growing importance of deep learning in effectively analyzing complex temporal data and improving decision-making across diverse applications [12, 29, 37, 17]. By integrating domain knowledge and leveraging specialized methodologies, these architectures significantly enhance the accuracy and applicability of imputation techniques across diverse applications.


(a) A neural network architecture with multiple layers and connections[39]


(b) Quantization and Activation Profiles of Various Neural Networks[44]



(c) Comparison of Test Accuracy Over Epochs for Different Data Sets and Non-IID Datasets[45]

Figure 5: Examples of Domain-specific and Application-driven Architectures

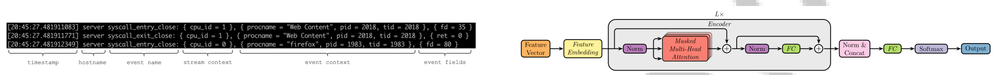
As shown in Figure 5, deep learning architectures have emerged as powerful tools for time series imputation, particularly when tailored to specific domains and applications. The figure showcases three distinct examples of such architectures, each highlighting a unique aspect of neural network design and application. The first example illustrates a neural network architecture characterized by its multi-layered structure and intricate connectivity, embodying the fundamental components of input, hidden, and output layers. This architecture underscores the importance of layer depth and interconnections in capturing complex temporal patterns. The second example delves into the quantization and activation profiles of various neural networks, offering insights into how different architectures manage precision and activation quantization. This highlights the critical role of precision in optimizing network performance for specific tasks. Lastly, the third example provides a comparative analysis of test accuracy across epochs for various datasets, including non-IID datasets, illustrating the adaptability and robustness of deep learning models in handling diverse data distributions. Together, these examples emphasize the necessity of domain-specific and application-driven approaches in developing effective deep learning architectures for time series imputation [39, 44, 45].

4 Imputation Techniques in Deep Learning

4.1 Traditional and Basic Imputation Techniques

Traditional imputation techniques, such as zero and mean imputation, provide a foundational approach to addressing missing data in time series analysis by substituting missing values with zero or the mean of available data. Despite their simplicity, these methods often fall short in capturing the intricate patterns of multivariate time series, leading to subpar predictive accuracy [37]. When juxtaposed with advanced deep learning models, which adeptly model complex temporal dependencies and variable interactions, the limitations of basic imputation become apparent. Methods like ARIMA, while common, struggle with missing data in non-stationary or volatile series [42]. In educational contexts, RNN-based models may inadequately capture the nuances of student learning processes, emphasizing the need for advanced techniques that integrate contextual information [6]. The absence of comprehensive guidelines for selecting suitable loss functions and metrics further complicates model development, highlighting the necessity for robust frameworks [14].

Basic imputation techniques, while foundational, are insufficient for the complex characteristics of multivariate time series, necessitating advanced deep learning methodologies. These sophisticated approaches, such as GP-VAE, enhance imputation accuracy and reliability by leveraging unsupervised feature learning and providing interpretable uncertainty estimates, thereby outperforming traditional methods in fields like healthcare and finance [29, 37].



(a) Server System Call Entries and Exit Logs[39] (b) Masked Multi-Head Attention Network[15]

Figure 6: Examples of Traditional and Basic Imputation Techniques

As depicted in Figure 6, deep learning addresses missing data through imputation techniques that aim to fill datasets' gaps, ensuring machine learning models' integrity and performance. Traditional methods estimate and replace missing values based on available information, while advanced approaches like the Masked Multi-Head Attention Network employ deep learning to effectively handle missing data by deriving meaningful representations even with incomplete data [39, 15].

4.2 Advanced Imputation Techniques

Advanced imputation techniques in deep learning significantly improve handling missing data in multivariate time series by employing sophisticated models that capture complex temporal dependencies. Variational autoencoders (VAEs) exemplify these techniques, offering a probabilistic framework that learns latent representations to generate plausible data points, enhancing data imputation through structured variational approximation and providing interpretable uncertainty estimates. Such advancements are crucial in complex domains like healthcare and finance, where missing values are prevalent and traditional techniques often fall short [15, 16, 1, 29, 28].

The MS-LSTM model, integrating various modules to predict intraday ranges based on historical data and economic patterns, showcases the ability of advanced methods to enhance imputation accuracy in financial time series. Attention networks, particularly weighted attention networks, further highlight these models' effectiveness, achieving superior hit ratios compared to traditional methods [23]. Innovative approaches like the Deep Convolutional Recurrent Autoencoder (DCRAE) combine convolutional neural networks for dimensionality reduction with recurrent neural networks for modeling dynamics, enabling efficient modeling of nonlinear systems. These techniques enhance data matching and imputation accuracy, significantly improving predictive performance across fields, including healthcare [15, 17, 1].

The K-UNN method, utilizing a tripled architecture informed by physical priors, exemplifies advanced techniques by improving reconstruction accuracy in MRI images, demonstrating the potential of domain-specific knowledge to enhance imputation outcomes. Incorporating differentiable data fusion, the Artificial Neural Twin (ANT) model allows backpropagation of loss gradients through distributed process steps, representing an advanced method that enhances model performance. The Associated Learning (AL) framework improves training efficiency and scalability, vital for advanced imputation techniques. The integration of synthetic examples into training datasets, particularly in recession forecasting models, leverages advancements in deep learning and generative adversarial networks (GANs) to tackle limited data availability for rare events. Utilizing models like DoppelGANger to generate high-fidelity synthetic time series data, researchers have shown that training on augmented datasets significantly improves predictive accuracy [46, 12, 19, 6, 4].

Exploring activation functions in neural networks suggests they enable modeling complex functions through infinite-dimensional polynomial representations, providing insights into developing more effective imputation strategies. The ongoing advancement of sophisticated techniques, such as the GP-VAE and NAIM model, underscores their importance in effectively addressing missing data issues in multivariate time series across diverse fields like healthcare and finance. These methods enhance the accuracy and reliability of time series analyses while providing interpretable uncertainty estimates and improving predictive performance without relying on traditional imputation approaches [15, 29].

4.3 Application-Specific Imputation Strategies

Application-specific imputation strategies address the unique challenges and requirements of various domains, enhancing the precision and applicability of imputation methods tailored to specific contexts. In healthcare, strategies manage missing physiological data by utilizing domain knowledge to improve the accuracy and reliability of medical diagnoses and treatment plans. These strategies often employ hybrid models combining convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to effectively capture both spatial and temporal dependencies in medical data, enhancing imputation outcomes by leveraging each model type's strengths [15, 47, 1, 30, 28].

In finance, imputation strategies navigate economic indicators and stock market data intricacies, where high volatility and long-range dependencies present significant challenges. Recent advancements, like transformer-based models such as "Not Another Imputation Method" (NAIM), offer innovative solutions by learning from available data without traditional techniques. This approach enhances predictive performance and resilience in the face of missing data, addressing the limitations of conventional methods that struggle with non-linear dynamics in financial datasets [46, 15, 1, 4]. Environmental monitoring applications significantly enhance accuracy by employing advanced imputation strategies that integrate sensor data with domain-specific features. These strategies leverage novel machine learning techniques, such as the transformer-based model NAIM, to effectively learn from incomplete datasets without traditional imputation [15, 48, 1, 40].

In transportation, imputation strategies address traffic flow and vehicle dynamics challenges. Domain-specific models often utilize RNNs and LSTMs to capture the sequential nature of traffic data, improving transportation networks' prediction and management. Implementing advanced strategies for traffic flow optimization and congestion reduction is essential for developing efficient transportation systems [15, 1, 40, 41, 43].

Application-specific imputation strategies represent a crucial advancement in deep learning for time series imputation, enabling models to effectively cater to various fields' distinct needs. By integrating domain-specific insights and employing advanced methodologies, these strategies significantly improve imputation techniques' accuracy and applicability across a wide range of applications [15, 1].

5 Applications and Case Studies

The integration of deep learning techniques has revolutionized various sectors by addressing the challenge of missing data through advanced imputation methods. This section highlights the diverse applications of deep learning imputation, focusing on its transformative impact across healthcare, finance, energy, transportation, and other domains, thereby enhancing data reliability and informing decision-making processes.

5.1 Healthcare Applications

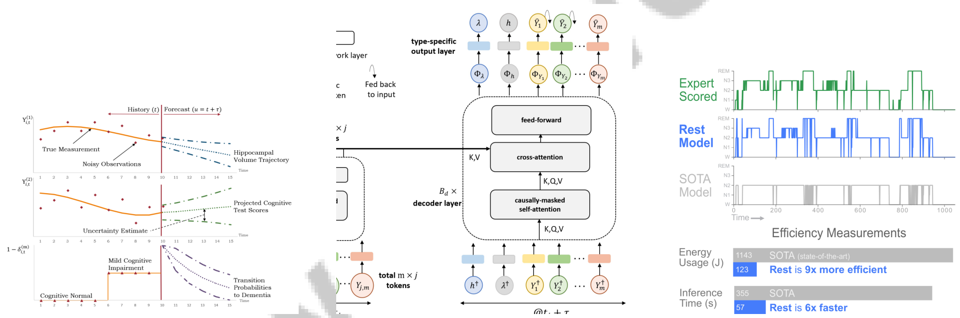
Advanced deep learning imputation techniques have markedly improved medical data analysis accuracy and reliability. For instance, the "Not Another Imputation Method" (NAIM) uses transformer-based models to manage missing values without traditional imputation, boosting predictive performance in incomplete datasets [15, 1]. These novel algorithms harmonize diverse health data sources, crucial for robust medical analyses, as incomplete datasets can severely impact diagnostics and treatment plans.

Experiments on datasets like Healing MNIST and SPRITES, alongside real-world data from the Physionet Challenge, underscore the efficacy of deep learning models in managing missing healthcare data [29]. These models leverage probabilistic frameworks to learn latent representations, generating plausible data points consistent with the underlying distribution, especially beneficial in high-dimensional medical datasets.

Models such as the ccRNN, tested on sensorimotor tasks, illustrate deep learning architectures' potential in real-world healthcare scenarios [21]. These models enhance prediction accuracy and facilitate personalized treatment plans by learning from incomplete data.

Incorporating domain-specific knowledge into deep learning models has proven essential for improving imputation outcomes. Innovative approaches harmonize data from diverse sources and utilize feature-specific embeddings to address missing values effectively, leading to significant improvements over existing techniques in both synthetic and real-world healthcare datasets [15, 1]. By integrating physiological signals and patient history, these models better capture the spatial and temporal dependencies inherent in medical data, resulting in more reliable imputation.

The use of deep learning imputation in healthcare data represents a significant advancement, offering accurate analyses that inform clinical decision-making and improve patient outcomes. For example, the transformer-based NAIM model effectively learns from incomplete data, while advanced feature matching methods integrate diverse medical data sources, leading to improved predictive performance and clinical insights [46, 15, 17, 1, 24].



(a) Modeling Cognitive Decline and Dementia[49]

(b) Deep Learning Model Architecture[28]

(c) Efficiency Comparison of Expert-Scoring, Rest Model, and SOTA Model in Sleep Monitoring[50]

Figure 7: Examples of Healthcare Applications

As illustrated in Figure 7, advanced modeling techniques and deep learning architectures are instrumental in addressing complex medical challenges. The "Modeling Cognitive Decline and Dementia" visual showcases a predictive model analyzing cognitive test scores, aiding early diagnosis and intervention planning. The "Deep Learning Model Architecture" image presents a recurrent neural network (RNN) design that enhances the model's ability to process complex healthcare data. Lastly, the "Efficiency Comparison of Expert-Scoring, Rest Model, and SOTA Model in Sleep Monitoring" image highlights the potential of innovative models to improve sleep analysis, contributing to better diagnosis and treatment of sleep disorders. Together, these examples underscore the transformative potential of technology in advancing healthcare solutions [49, 28, 50].

5.2 Finance and Economic Forecasting

Deep learning applications in finance and economic forecasting have significantly enhanced prediction accuracy, particularly in stock market analysis and economic indicators. These models capture complex temporal dependencies and nonlinear relationships that traditional methods often overlook, facilitating robust trading strategies and accurate economic trend forecasts. Recent studies reveal that deep learning not only improves prediction accuracy but also extracts meaningful patterns from vast datasets, aiding informed decision-making in finance [12, 37, 46].

A notable application involves predicting stock market movements following financial disclosures, leveraging deep learning's capacity to process unstructured data and extract relevant patterns [46]. Advanced neural network architectures analyze the impact of financial news on stock prices, providing valuable insights for investors.

The LSTM-GRU model exemplifies deep learning's effectiveness in financial forecasting, outperforming traditional methods in stock price predictions and trading strategy development [51]. Its ability to capture long-term dependencies and adapt to market dynamics highlights its superiority over conventional approaches.

Transformer architectures have proven advantageous in financial contexts due to their capability to manage long-range dependencies and inherent volatility in economic time series data, enhancing predictive accuracy and facilitating sophisticated analyses of financial trends [12, 23]. These models utilize attention mechanisms to prioritize relevant signals, improving the robustness of imputation tasks in finance.

The integration of deep learning techniques marks a transformative shift in finance, as models like attention LSTM and convolutional neural networks demonstrate superior predictive capabilities compared to traditional methods. They effectively capture complex patterns in financial data, enhancing decision analytics and enabling practitioners to make informed decisions in a dynamic market environment [46, 17, 12, 23, 4]. As these models evolve, they hold the potential to revolutionize financial analytics, providing more accurate predictions that inform investment decisions and economic policies.

5.3 Energy and Environmental Monitoring

The integration of deep learning technologies into energy and environmental monitoring has become essential, driven by the need for precise data analysis. Recent advancements have shown significant improvements in energy time-series forecasting, including electric load demand and renewable energy generation predictions, enhancing data interpretation and decision-making in response to energy consumption and sustainability challenges [46, 12, 33, 4]. Deep learning models effectively capture complex temporal patterns in time series data, crucial for monitoring and forecasting in these sectors.

In energy, combining time-series forecasting with deep reinforcement learning has shown promise in optimizing energy arbitrage strategies. Utilizing multiple forecasts alongside deep reinforcement learning can lead to a 60

Environmental monitoring benefits from deep learning models that process large volumes of sensor data to track climate patterns. These models effectively handle incomplete datasets, enhancing predictive accuracy and operational performance [46, 15, 17, 6, 4]. By leveraging neural networks to model spatial and temporal dependencies, deep learning techniques improve environmental data imputation and forecasting, facilitating informed management and policy-making.

The application of deep learning in energy and environmental monitoring represents a significant advancement, providing robust solutions to the challenges presented by complex datasets. As these techniques continue to develop, they can enhance energy management and conservation initiatives, supporting sustainable development goals through improved predictive models and decision-making capabilities [46, 33, 52, 53, 4].

5.4 Transportation and Traffic Analysis

Deep learning applications in transportation and traffic analysis have significantly improved traffic flow prediction and system optimization. The DeepTFP model exemplifies advancements in this domain, enhancing prediction accuracy by analyzing complex traffic data patterns [41].

Beyond traffic flow prediction, deep learning models analyze travel time and speed data, essential for predicting travel times within networks. Comparative analyses highlight the importance of these datasets in enhancing prediction precision, improving transportation system efficiency [42]. By incorporating these data types, deep learning models better capture dynamic traffic patterns, leading to informed decision-making and optimized traffic control strategies.

The incorporation of deep learning techniques marks a substantial leap in transportation analysis, enabling sophisticated predictive models that address the complexities of dynamic traffic datasets. For instance, the DeepTFP algorithm utilizes deep residual neural networks to analyze mobile data, allowing accurate traffic flow predictions by considering various influencing factors. Additionally, innovative approaches integrating time-series data with unstructured textual information have improved accuracy in forecasting taxi demand in event areas, showcasing deep learning's potential to enhance transportation planning and congestion management [43, 41]. As these models evolve, they hold the potential to transform transportation systems, supporting the development of smarter urban mobility solutions.

5.5 Other Applications

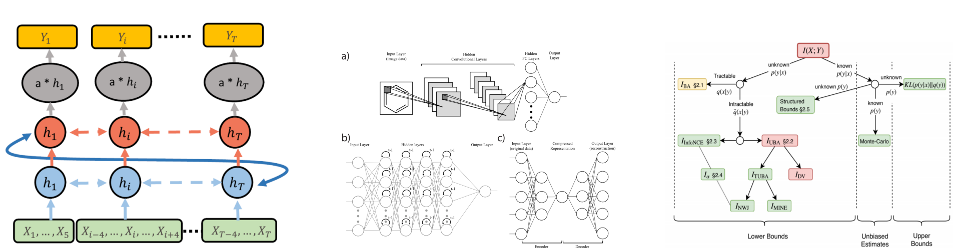
Deep learning imputation techniques extend beyond traditional domains, showcasing versatility in addressing missing data challenges across various fields. In sports analytics, deep learning models reconstruct incomplete player performance datasets, significantly improving performance evaluations and strategic decision-making [46, 17, 4]. By leveraging temporal dependencies captured by architectures like RNNs and LSTMs, these models facilitate the imputation of missing performance metrics, providing coaches and analysts with comprehensive insights into player development.

In agriculture, deep learning techniques have been applied to impute missing data in crop monitoring systems, where environmental and sensor data are often incomplete. Hybrid architectures combining CNNs and RNNs effectively capture spatial and temporal patterns, enhancing the accuracy of reconstructing missing data points essential for reliable crop yield predictions [10, 31]. This application optimizes resource allocation and improves food security by providing farmers with reliable data for decision-making.

The entertainment industry has enhanced content recommendation systems through advanced deep learning imputation methods, such as transformer-based models that handle missing data effectively. These methods utilize feature-specific embeddings and masked self-attention mechanisms to improve predictive performance, enabling personalized recommendations and enhancing user engagement [15, 1, 17, 40]. By addressing missing user preference data, deep learning models refine recommendation algorithms, optimizing user experience.

In telecommunications, deep learning models impute missing data in network traffic analysis, where incomplete logs hinder performance monitoring. Utilizing Transformer models and self-attention mechanisms, these techniques enable precise reconstruction of missing traffic data, enhancing predictive analytics and optimizing resource allocation strategies. Evaluations against traditional imputation techniques demonstrate superior performance in handling incomplete datasets [15, 1, 17, 12, 29].

The widespread implementation of deep learning imputation techniques across fields such as healthcare, natural language processing, and bioinformatics highlights their potential to effectively tackle missing data challenges. These advanced methods enhance the accuracy and reliability of analyses and predictions, offering innovative solutions that surpass traditional techniques, as demonstrated by the transformer-based NAIM and novel algorithms for schema matching and data harmonization [13, 15, 1, 17]. As these models evolve, they promise to transform industries by providing robust solutions to data incompleteness, supporting informed decision-making and strategic planning.



(a) Deep Learning Model with Multiple Layers and Connections[54]

(b) Convolutional Neural Networks (CNNs) and Deep Learning Models[55]

(c) Flowchart Illustrating Methods for Estimating Mutual Information[56]

Figure 8: Examples of Other Applications

As shown in Figure 8, the diverse applications and case studies of deep learning are visually represented, highlighting various models and methodologies. The first example illustrates a deep learning model with multiple layers, showcasing its architecture's capacity to process diverse datasets. The second example focuses on Convolutional Neural Networks (CNNs), detailing their architecture and functionality in image recognition tasks. Lastly, the flowchart depicts methods for estimating mutual information between random variables, providing insight into the statistical techniques employed in machine learning applications. Together, these examples underscore the technical intricacies of deep learning models and their profound impact across various fields [54, 55, 56].

6 Challenges and Future Directions

Exploring the challenges of deploying deep learning models reveals critical issues affecting their effectiveness, notably scalability and computational efficiency. These challenges are particularly pronounced in dynamic environments requiring real-time processing. The following subsections delve into these challenges, underscoring the necessity for advancements in model optimization and resource management to facilitate the broader adoption of deep learning across various domains.

6.1 Scalability and Computational Efficiency

Scalability and computational efficiency are crucial challenges in deploying deep learning models, especially with high-dimensional datasets. The intensive computational demands hinder scalability, as seen in financial markets where optimal currency pair selection and lagging predictions require substantial resources [18]. These issues are compounded by reliance on historical data, which may not accurately reflect future dynamics, particularly in volatile markets [23]. The need for large datasets and high computational costs poses significant barriers to the effective application of deep learning [13]. Bayesian models, for instance, require efficient integration of robust priors for scalability [11]. The Artificial Neural Twin (ANT) framework addresses these challenges by enabling continual adaptation, enhancing resource utilization [2]. The Associated Learning (AL) framework improves scalability by allowing simultaneous updates across layers, alleviating computational bottlenecks [5]. Nonetheless, challenges remain in domains needing real-time analysis, where computational efficiency is critical. Addressing these issues is essential for the broader applicability of deep learning models, necessitating ongoing research in model optimization and resource management to enhance predictive performance and operational efficiency [13, 46, 17, 40, 4].

6.2 Interpretability and Model Transparency

Interpretability and transparency of deep learning models are critical challenges that hinder their adoption. Often viewed as 'black boxes,' these models obscure decision-making processes, limiting applicability in real-world settings [13]. This lack of transparency is concerning in sensitive areas such as healthcare, finance, and energy systems, where understanding model predictions is vital for reliability and accountability. In healthcare, model opacity can impede integration into clinical workflows, where clear explanations are necessary for patient care decisions. In finance, interpreting model predictions is crucial for robust trading strategies and market insights. Recent advancements, including attention-based models and LSTMs, have enhanced predictive performance and provided valuable visualizations for understanding market trends [46, 23, 4, 18]. The complexity of deep learning models can also introduce biases, affecting fairness and accuracy. Future research should focus on methodologies to enhance interpretability, demystifying internal workings and providing actionable insights. Improved interpretability increases transparency and trust, making models more effective in supporting decision-making across diverse fields such as finance, business analytics, and healthcare, facilitating better understanding and communication of predictions [46, 13, 4].

6.3 Adaptive and Robust Model Development

Advancing adaptive and robust model development is crucial for enhancing deep learning techniques in time series imputation, particularly for multivariate datasets. Future research should optimize existing architectures and explore innovative methodologies to enhance adaptability and robustness. Refining architectures such as the Deep Convolutional Recurrent Autoencoder (DCRAE) through diverse training strategies and broader application could validate its effectiveness further [20]. In-

AI-generated, for reference only.

corporating additional calibration data and exploring inter-frame similarities, as seen in the K-UNN method, are potential avenues for more adaptive models [3]. Integrating diverse datasets and hybrid models combining traditional and deep learning approaches could enhance predictive capabilities, particularly in financial forecasting. Establishing standardized guidelines for selecting loss functions and metrics ensures consistency and effectiveness in model evaluation [14]. Exploring novel architectures and enhancing interpretability are vital for improving applicability in real-world scenarios [13]. This includes refining hyperparameter selection and examining additional association measures for feature matching, which can bolster model adaptability and robustness [1]. Future research could also expand linear algebra operations as differentiable operators, enhancing adaptability and robustness [57]. Optimizing network architectures through evolutionary strategies presents another promising research direction [31]. By focusing on these areas, future advancements in deep learning can yield more sophisticated and reliable imputation techniques applicable across a wide array of domains.

7 Conclusion

This survey highlights the profound influence of deep learning on the imputation of multivariate time series, underscoring its pivotal role in advancing data analysis and prediction across various fields. The implementation of deep neural networks has revolutionized automatic feature extraction, significantly boosting predictive capabilities in sectors such as computational chemistry. Models like Disease-Atlas exemplify substantial performance improvements over traditional approaches, showcasing their potential in clinical decision-making, particularly in managing conditions like Alzheimer's Disease.

The GP-VAE model further demonstrates the superiority of deep learning techniques, outperforming classical methods in imputation quality while providing interpretable uncertainty estimates crucial for informed decision-making. These advancements affirm deep learning's critical function in addressing the challenges associated with missing data, thereby enhancing the precision and dependability of time series analyses.

Additionally, this survey positions deep learning as a vital tool for big data analysis, emphasizing the necessity for continued research to refine existing methodologies and address current limitations. The prospects for future advancements in this field are considerable, with the potential to develop more adaptive and resilient models capable of navigating the complexities of high-dimensional datasets and dynamic environments.

References

[1] Sandhya Tripathi, Bradley A. Fritz, Mohamed Abdelhack, Michael S. Avidan, Yixin Chen, and Christopher R. King. Deep learning to jointly schema match, impute, and transform databases, 2022.

[2] Johannes Emmert, Ronald Mendez, Houman Mirzaalian Dastjerdi, Christopher Syben, and Andreas Maier. The artificial neural twin – process optimization and continual learning in distributed process chains, 2024.

[3] Zhuo-Xu Cui, Sen Jia, Qingyong Zhu, Congcong Liu, Zhilang Qiu, Yuanyuan Liu, Jing Cheng, Haifeng Wang, Yanjie Zhu, and Dong Liang. K-unn: k-space interpolation with untrained neural network, 2022.

[4] Mathias Kraus, Stefan Feuerriegel, and Asil Oztekin. Deep learning in business analytics and operations research: Models, applications and managerial implications, 2019.

[5] Yu-Wei Kao and Hung-Hsuan Chen. Associated learning: Decomposing end-to-end backpropagation based on auto-encoders and target propagation, 2021.

[6] Shalini Pandey, George Karypis, and Jaideep Srivastava. An empirical comparison of deep learning models for knowledge tracing on large-scale dataset, 2021.

[7] Maciej Śliwowski, Matthieu Martin, Antoine Souloumiac, Pierre Blanchart, and Tetiana Aksenova. Impact of dataset size and long-term ecog-based bci usage on deep learning decoders performance, 2022.

[8] John Chiang. Activation functions not to active: A plausible theory on interpreting neural networks, 2023.

[9] Xiangyi Li and Huaming Wu. Spatio-temporal representation with deep neural recurrent network in mimo csi feedback, 2019.

[10] Si-An Chen, Chun-Liang Li, Nate Yoder, Sercan O. Arik, and Tomas Pfister. Tsmixer: An all-mlp architecture for time series forecasting, 2023.

[11] Daniel T. Chang. Probabilistic deep learning with probabilistic neural networks and deep probabilistic models, 2021.

[12] Saptarshi Sengupta, Sanchita Basak, Pallabi Saikia, Sayak Paul, Vasilios Tsalavoutis, Frederick Atiah, Vadlamani Ravi, and Alan Peters. A review of deep learning with special emphasis on architectures, applications and recent trends, 2019.

[13] Mohammad Mustafa Taye. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers*, 12(5):91, 2023.

[14] Juan Terven, Diana M. Cordova-Esparza, Alfonso Ramirez-Pedraza, Edgar A. Chavez-Urbiola, and Julio A. Romero-Gonzalez. Loss functions and metrics in deep learning, 2024.

[15] Camillo Maria Caruso, Paolo Soda, and Valerio Guarrasi. Not another imputation method: A transformer-based model for missing values in tabular datasets, 2024.

[16] Heejoon Koo and To Eun Kim. A comprehensive survey on generative diffusion models for structured data, 2023.

[17] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and Sundaraja S Iyengar. A survey on deep learning: Algorithms, techniques, and applications. *ACM computing surveys (CSUR)*, 51(5):1–36, 2018.

[18] Shujian Liao, Jian Chen, and Hao Ni. Forex trading volatility prediction using neural network models, 2021.

[19] Sam Dannels. Creating disasters: Recession forecasting with gan-generated synthetic time series data, 2023.

[20] Francisco J. Gonzalez and Maciej Balajewicz. Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems, 2018.

[21] Joseph Pemberton, Ellen Boven, Richard Apps, and Rui Ponte Costa. Cortico-cerebellar networks as decoupling neural interfaces, 2021.

[22] Yi-Hsiu Liao, Hung-Yi Lee, and Lin shan Lee. Towards structured deep neural network for automatic speech recognition, 2015.

[23] Sangyeon Kim and Myungjoo Kang. Financial series prediction using attention lstm, 2019.

[24] Soham Gadgil, Joshua Galanter, and Mohammadreza Negahdar. Transformer-based time-series biomarker discovery for copd diagnosis, 2024.

[25] Vivien Sainte Fare Garnot, Lynsay Spafford, Jelle Lever, Christian Sigg, Barbara Pietragalla, Yann Vitasse, Arthur Gessler, and Jan Dirk Wegner. Deep learning meets tree phenology modeling: Phenoformer vs. process-based models, 2024.

[26] Chang Gao, Shu-Fu Shih, J. Paul Finn, and Xiaodong Zhong. A projection-based k-space transformer network for undersampled radial mri reconstruction with limited training subjects, 2022.

[27] Haoran Zhao and Wayne Isaac Tan Uy. Genformer: A deep-learning-based approach for generating multivariate stochastic processes, 2024.

[28] Zhiyue Zhang, Yao Zhao, and Yanxun Xu. Transformerlsr: Attentive joint model of longitudinal data, survival, and recurrent events with concurrent latent structure, 2024.

[29] Vincent Fortuin, Dmitry Baranchuk, Gunnar Rätsch, and Stephan Mandt. Gp-vae: Deep probabilistic time series imputation, 2020.

[30] Linpeng Jin. Go beyond multiple instance neural networks: Deep-learning models based on local pattern aggregation, 2022.

[31] Aditya Rawal and Risto Miikkulainen. From nodes to networks: Evolving recurrent neural networks, 2018.

[32] Julie Keisler, Sandra Claudel, Gilles Cabriel, and Margaux Brégère. Automated deep learning for load forecasting, 2024.

[33] Maria Tzelepi, Charalampos Symeonidis, Paraskevi Nousi, Efstratios Kakaletsis, Theodoros Manousis, Pavlos Tosidis, Nikos Nikolaidis, and Anastasios Tefas. Deep learning for energy time-series analysis and forecasting, 2023.

[34] Aditya Datar and Pramit Saha. The promise of analog deep learning: Recent advances, challenges and opportunities, 2024.

[35] Alireza Morsali, Afshin Haghighat, and Benoit Champagne. Deep learning framework for hybrid analog-digital signal processing in mmwave massive-mimo systems, 2021.

[36] Gabriela Gómez Jiménez and Demian Wassermann. Deep multivariate autoencoder for capturing complexity in brain structure and behaviour relationships, 2024.

[37] John Cristian Borges Gamboa. Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*, 2017.

[38] Yongxin Liu, Jian Wang, Jianqiang Li, Shuteng Niu, and Houbing Song. Zero-bias deep learning enabled quick and reliable abnormality detection in iot, 2021.

[39] Quentin Fournier, Daniel Aloise, Seyed Vahid Azhari, and François Tetreault. On improving deep learning trace analysis with system call arguments, 2021.

[40] Jingzhi Gong. Pushing the boundary: Specialising deep configuration performance learning, 2025.

[41] Yuanfang Chen, Falin Chen, Yizhi Ren, Ting Wu, and Ye Yao. Deeptfp: Mobile time series data analytics based traffic flow prediction, 2017.

[42] Armstrong Aboah and Elizabeth Arthur. Comparative analysis of machine learning models for predicting travel time, 2021.

[43] Filipe Rodrigues, Ioulia Markou, and Francisco Pereira. Combining time-series and textual data for taxi demand prediction in event areas: a deep learning approach, 2018.

[44] Alberto Delmas, Sayeh Sharify, Patrick Judd, Kevin Siu, Milos Nikolic, and Andreas Moshovos. Dpred: Making typical activation and weight values matter in deep learning computing, 2018.

[45] Sahil Tyagi and Martin Swany. Scadles: Scalable deep learning over streaming data at the edge, 2024.

[46] Stefan Feuerriegel and Ralph Fehrer. Improving decision analytics with deep learning: The case of financial disclosures, 2018.

[47] Louis Falissard, Karim Bounebache, and Grégoire Rey. Learning a binary search with a recurrent neural network. a novel approach to ordinal regression analysis, 2021.

[48] Nasser Kazemi. Across-domains transferability of deep-red in de-noising and compressive sensing recovery of seismic data, 2020.

[49] Bryan Lim and Mihaela van der Schaar. Forecasting disease trajectories in alzheimer's disease using deep learning, 2018.

[50] Rahul Duggal, Scott Freitas, Cao Xiao, Duen Horng Chau, and Jimeng Sun. Rest: Robust and efficient neural networks for sleep monitoring in the wild, 2020.

[51] Qishuo Cheng. Intelligent optimization of mine environmental damage assessment and repair strategies based on deep learning, 2024.

[52] Sarvesh Patil. Deep learning based natural language processing for end to end speech translation, 2018.

[53] Mulomba Mukendi Christian, Yun Seon Kim, Hyebong Choi, Jaeyoung Lee, and SongHee You. Enhancing wind speed and wind power forecasting using shape-wise feature engineering: A novel approach for improved accuracy and robustness, 2024.

[54] E. A. Huerta and Zhizhen Zhao. Advances in machine and deep learning for modeling and real-time detection of multi-messenger sources, 2021.

[55] Garrett B Goh, Nathan O Hodas, and Abhinav Vishnu. Deep learning for computational chemistry. *Journal of computational chemistry*, 38(16):1291–1307, 2017.

[56] Marshal Arijona Sinaga. On study of mutual information and its estimation methods, 2021.

[57] Matthias Seeger, Asmus Hetzel, Zhenwen Dai, Eric Meissner, and Neil D. Lawrence. Auto-differentiating linear algebra, 2019.