International Journal of Applied Research 2025; 11(2): 101-107

International Journal of Applied Research

ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor (RJIF): 8.4
IJAR 2025; 11(2): 101-107
www.allresearchjournal.com
Received: 05-02-2025
Accepted: 19-02-2025

Zainab Qahtan Mohammed
University of Diyala, College of Basic education, Iraq

Corresponding Author:
Zainab Qahtan Mohammed
University of Diyala, College of Basic education, Iraq

Theoretical insights into deep learning models in high-dimensional feature spaces: Performance and stability analysis

Zainab Qahtan Mohammed

DOI: https://dx.doi.org/10.22271/allresearch.2025.v11.i2b.12343

Abstract
This research investigates the delving parameters of the theoretical deep learning models in high-dimensional feature spaces and their delicate balance between performance and stability. The study starts with a comprehensive review of extant scholarship covering the action of deep learning in hugely populated settings. Then a theoretical approach explaining the influence of high-dimensionality on the behavioral aspects of models, especially in regard to the proposed ratio of stability and performance, is provided. Moreover, these approaches are theoretical and mathematical with the aim of providing the model stability under these conditions. The study also searches phenomenal extensions of computer vision and data modeling using the principles learned in this article.

Keywords: Deep learning models, high-dimensional feature spaces, performance and stability, theoretical deep learning

1. Introduction

1.1 Background on Deep Learning Models

Deep learning works in on the premise of automating the extraction of meaningful information from large data sets by classifying them. It was built on traditional artificial neural networks which gave the ability to build sophisticated deep learning models that learn from high dimensional data. Such models are usually composed of an input layer, hidden layers, and an output layer through which the model learns to discriminate features without being heavily supervised.

More recently, extensive research efforts have been undertaken in deep learning, especially for image and video analysis, and natural language processing due to increased computational capabilities and data availability. Deep neural networks, DNNs, are very effective in extracting features from data of high dimension due to the existence of DNNs, which solve complex tasks that are not possible with other methods [1]. The current deep learning architectures are improved with methodologies of transfer learning and pre-trained weights, that enable the models trained on large datasets to be fine-tuned to particular problems with very little additional data, reducing computation and maintaining efficacy [4].

DNN embedding techniques are also more efficient as they work by mapping high dimensional inputs into a lower dimension without losing information which increases speed in training and inference [7].

Understanding how deep learning navigates high-dimensional feature spaces is crucial for both research and practical applications, such as medical diagnostics and image recognition [3]. Researchers aim to optimize architectural choices to improve model stability and generalization across various fields.

1.2. The Significance of High-Dimensional Spaces

The efficiency of deep learning models is significantly enhanced by high-dimensional spaces, particularly for complex datasets with many attributes. The 'curse of dimensionality' describes issues like data sparsity, overfitting, and computation inefficiency which are especially relevant with high dimensional data. Addressing these problems is particularly important in areas like image classification, which is low hanging proving for AI/ML tools since datasets contain thousands of dimensions [7]. Removing some dimensions while preserving crucial information is an approach known as embedding which in turn aids in uncovering new patterns. These algorithms are very useful in identifying important relationships in dense higher dimensional spaces.

Techniques like UMAP and Locally Linear Embedding (LLE) seem to excel at reducing dimensions while maintaining important local structures within the broader context of the data metrics [2]. This is crucial for applications that involve non-linear interactions between different features.

Model performance rapidly increases with the understanding the relationships and geometric structures of higher dimensional spaces. For instance, designing machine learning interatomic potentials requires navigating multiple properties simultaneously, highlighting both the advantages and challenges of high-dimensional analysis. Recognizing the importance of high-dimensional spaces fosters the development of innovative methods for handling complex datasets, while addressing the challenges that come with increased dimensions.

1.3. Objectives of the Research

This research aims to identify and solve the problems encountered when developing deep learning techniques systems in high-dimensional feature spaces. We wish to construct a comprehensive automated assessing system that highlights critical factors affecting model performance within these complex environments. Specifically, we will look at the influence of different sample sizes on accuracy and stability as are the concerns is so pervasive in many of deep learning applications. In addition, we will examine the relationship of high-dimensional spaces with the processes of learning and strive to provide some explanations to phenomena that are observed.

Another key objective is to perform deep learning model comparisons for high dimensional context with lower dimensional context and the corresponding literature gap on these issues [4]. This study will provide insight on classification performance variances that results from differences in dimensionality that have conjectured on before. Also, we attempt to fill the gaps that have been left existing by integrating other diverse fundamental approaches in particular fields like generalization, vision, medical diagnosis and class imbalance factors.

By pursuing these goals, we aim to offer valuable insights that can help practitioners choose the most suitable models for their specific applications, while also establishing a foundation for future research efforts focused on refining deep learning techniques to achieve better performance across various scenarios.

2. Literature Review

2.1. Overview of Existing Research on Deep Learning

The advancement of deep learning sparked marked interest in it's study because of the ability to formulate and capture intricate relationships in multidimensional space. Humaidi *et al*. mention the importance of convolutional neural networks (CNNs) that are often employed for many challenges especially in image processing and even in biomedical fields, but also data imbalance, low interpretability, and overfitting.

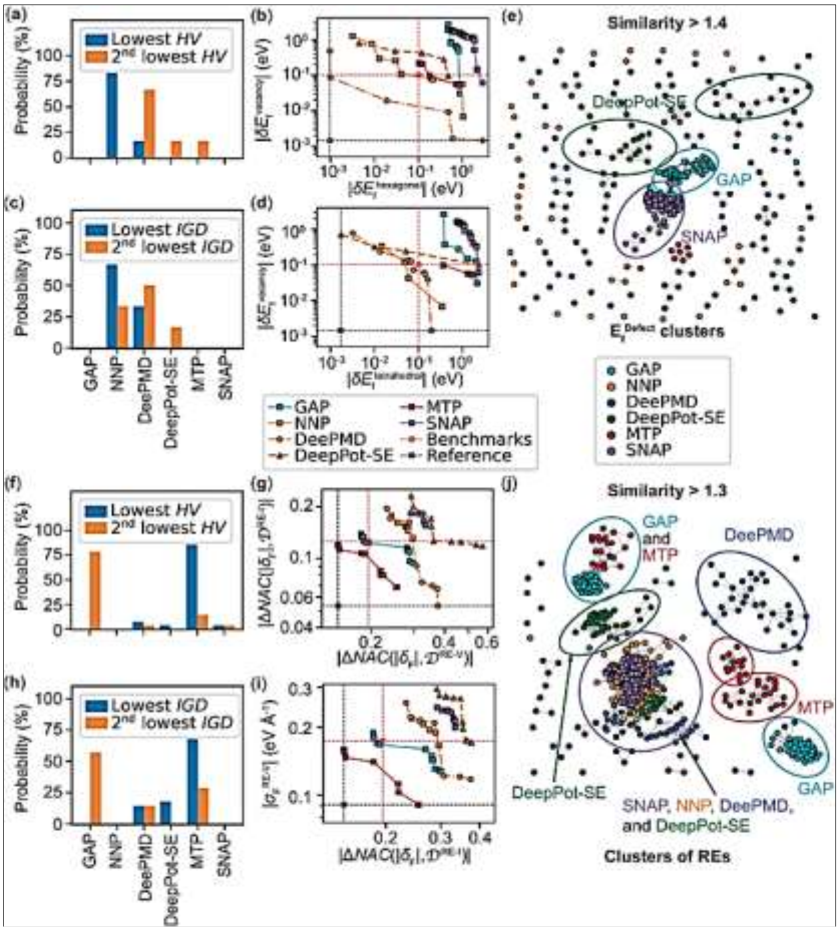

Fig 1: The fraction of the optimal MLIP models for each type of MLIPs with the lowest and the 2 nd lowest a HV.

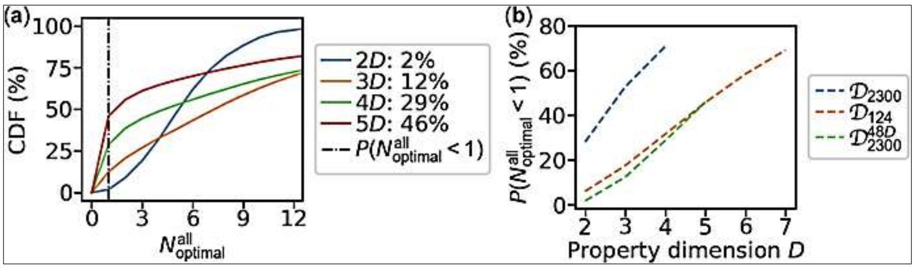

Fig 2: The joint performances of MLIPs for all combinations of properties

According to Liu *et al*., deep learning applications already outperformed the classic approaches in computer vision because it is based on hierarchical neural networks that tremendously enhance feature extraction and semantic understanding in tasks such as image classification and even object detection. On the other hand, try to address the problem of labeled dataset scarcity for model performance where they call for more exploration of building smaller training set or synthetic data to improve the model's performance.

The benefits of deep learning are evident in the agricultural Industry when it comes to weed detection and precision management. Ahmad's study suggests that trained models can separate weeds from crops very well when enough data is made available for training purposes.

Additionally, Chhabra *et al*. highlight the importance of identifying harmful training samples that could negatively impact model performance, tackling the ongoing challenges of maintaining high-quality datasets during training. These studies demonstrate a growing focus on improving deep learning techniques while addressing issues related to scalability and interpretability across different domains.

2.2. Stability and Performance in High-Dimensional Contexts

In the realm of high-dimensional data, deep learning models grapple with considerable challenges that can hinder their performance and reliability. One prominent concern is the 'curse of dimensionality,' which encompasses various issues that arise when examining data in vast feature spaces. With the progression in the number of dimensions, the obtuse sparsity results in spread separation of data points, thereby thwarting clustering and classification attempts. The increased sparsity creates the need for a much larger volume of training data to obtain reliable results which unfortunately expands at an exponential rate compared to the dimensions involved.

Moreover, overfitting is further exacerbated in higher dimensions because models tend to capture noise instead of relevant patterns. This problem arises from a relatively complex model that is able to incorporate spurious relationships that are made but fail to generalize to new data, as pointed out in (1). Even more puzzling is the fact that traditional distance measures, such as the Euclidean distance, become irrelevant in high dimensions, making it even more difficult for an algorithm to measure relationships between data points with precision.

New approaches focused on improving model interpretability and stability are emerging, often through regularization techniques or innovative changes to model structures. One such approach involves using Fourier-transform-based attribution priors, which help enhance sensitivity to random initialization without losing interpretability.

Additionally, research on parallel and sequential architectures in deep reinforcement learning is gaining attention. However, these approaches introduce their own set of challenges, such as coordination issues and complications related to gradients.

These complexities underscore the urgent need to develop specialized strategies that can effectively address the unique challenges of high-dimensional data environments, all while ensuring stable and reliable performance across a range of applications.

2.3. Gaps in Current Knowledge

In the realm of deep learning lies unsolved puzzles that are magnified when the issue of high-dimensional feature spaces or model performance are considered. One of the major problems is the generalization inconsistency across multiple datasets. For example, models that have been trained on one dataset tend to perform poorly on a different one because of completely different data and uncontrolled factors outside the training scope. This puts the durability and portability of deep learning models into question, especially in medical imaging.

While there is hope of doing something in this area given the recent focus on hyperparameters and their effects on performance, there really are no clear rules on what type of data is needed. The direct link between accuracy of the model and the sample size per class is intricate and deeply ignored, which makes the problem of collecting data for real world scenarios tougher than it seems.

Yet another aspect that has not been thoroughly researched is the impact of concept drift— changes in distribution of data over a period of time—on the models. Studies on how to make models persistently adjustable or relearn over a period of time are rather scarce.

Moreover, while many studies concentrate on empirical results from deep learning applications, there is a lack of theoretical insights to back up these practices. Understanding why certain techniques perform well in high-dimensional environments is critical for both academic research and real-world applications [6]. Bridging the gap between theory and practice is essential for the advancement of future research.

3. Theoretical Framework

3.1. Development of a Theoretical Model

Gaining an insight into estimating the performance of deep learning methods in high-dimensional feature spaces needs the development of a theoretical framework. The first step would be to analyze the interaction between deep learning models and high-dimensional datasets. In particular, the study needs to focus on diffusion models as unsupervised learners where their convergence sampling and statistical features are learned and further needed [6]. Insights into how score functions are learned and utilized is one such key area that is heavily dependent on the statistical features that can negate the curse of dimensionality.

When guidance mechanisms are incorporated into conditional diffusion models, it adds a new level of complexity, thus, requiring the development of new training paradigms where conditional scores can be estimated and linked to unconditional scores [6]. Performance on complex input models can greatly benefit from adaptive techniques.

Deep learning and deep reinforcement learning along with their corresponding architectures have profound impacts on neural representation of pompous input data. Understanding how deep reinforcement learning tackles state-space abstraction can lead to theoretical breakthroughs in deep learning models, particularly in real-world applications with continuous, high-dimensional inputs.

In conclusion, creating a comprehensive theoretical model necessitates integrating ideas from various research fields, while considering empirical findings and mathematical formulations related to statistical estimation and performance improvement.

3.2. Interaction between High-Dimensional Spaces and Learning Dynamics

The vast characteristic space offers certain difficulties for deep learning models because of the 'curse of dimensionality' and data sparsity experienced reduces model performance. The sparsity of data results in poor categorization and makes it difficult to detect useful patterns.

Furthermore, the sparsely populated high level feature space and learning processes can produce disturbing consequences. For instance, a low dimensional abstraction is created internally and evidence shows that agents in some deep RL frameworks can be trained with high-dimensional inputs. Hence, supercomplex environments may existe where one models are able to develop sophisticated internal representations that would improve decision making.

The design of neural networks is the single most important factor for dealing with high dimensional datasets. Locally linear embeddings (LLE) attempt to unravel the structure of high-dimensional data within lower-dimensional space while preserving the local geometric features around neighbourhoods [2]. This method attempts to solve one of the problems that arise with high dimensionality.

In addition, nearly all diffusion models are known as being perfect at capturing high-dimensional data due to their power in sampling in controlled conditions [6].

By approaching optimization challenges as sampling tasks, these models show great potential in handling the complexities that arise from high dimensionality. Understanding how deep learning models interact with high-dimensional spaces is crucial for improving their performance and stability, paving the way for future advancements in methodologies across a wide range of applications.

4. Performance Analysis of Deep Learning Models in High Dimensions

4.1. Factors Influencing Model Performance

The effectiveness of deep learning models is significantly impacted by several factors within high-dimensional feature spaces. One key factor is the selection of input features, which can be optimized using dimensionality reduction techniques like Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE). These methods reduce the data's dimensionality while keeping its most important components, helping to minimize the risk of overfitting and reducing computational strain [2].

Moreover, feature selection is imperative as it determines a suitable portion of features that would best toughen a model's ability to train and generalize. For instance, too many features may increase noise or redundancy, which complicate the learning process, and in turn, decrease predictive performance. Additionally, correlated features may inhibit interpretability through redundantly supplying the same information.

The model's efficiency would also be reliant on the specific deep learning architecture selected, as different structures have the ability to capture varying non-linearities embedded in high-dimensional data. For example, omics data are often hierarchical and complex, hence V-E Autoencoders Change Omics Data to Lower Dimensional Latent Spaces to Capture These Complicated Patterns [3].

Another critical aspect of performance would be hyperparameter tuning. As one of the most sensitive factors, hyperparameters may yield diverse results across varying datasets if not changed appropriately.

Additionally, larger training datasets typically enhance classification accuracy, though the relationship between sample size and performance can vary depending on the specific task and the characteristics of the dataset.

4.2. Comparative Studies with Low-Dimensional Models

There are marked differences in performance between deep learning models and traditional low-dimensional methods in most comparative studies. The deep learning model also proved capable of identifying complex patterns in high-dimensional data with an R^2 score of 0.996, outperforming random forests (0.886), XGBoost (0.927), and ANN (0.917) in groundwater quality prediction.

Deep learning approaches, such as variational autoencoders, have outperformed classical algorithms in high-dimensional scenarios like cancer classification with omics data by capturing intricate non-linear interactions that traditional methods often miss [3]. This advantage becomes particularly evident when dimensionality exceeds sample size, where simpler models are more prone to overfitting.

On a global scale, deep learning frameworks have demonstrated significantly greater adaptability and accuracy in multivariate probabilistic energy forecasting across diverse datasets compared to low-dimensional methods. While low-dimensional models are easier to interpret and train, they often struggle to capture complex interrelations within large or deep feature sets.

However, challenges remain in hyperparameter tuning and computational efficiency—higher dimensionality increases the difficulty of model training and deployment. Although low-dimensional methods can be useful for tasks requiring interpretability or when data is limited, deep learning remains the superior choice in high-dimensional contexts, where it can extract deeper insights from data.

5. Stability Analysis under High-Dimensional Conditions

5.1. Mathematical Techniques for Stability Assessment

A range of mathematical approaches is used to assess the robustness of deep learning frameworks. One widely recognized technique is Aggressive Stochastic Weight Averaging (ASWA), which fine-tunes model weights to minimize the impact of random seed variations on stability. Changes in performance metrics across various attempts are influenced by random seeds. Averaging weights from different seeds will yield a result that reduces the standard deviation in metrics by 72% as noted in [9]. In addition, some steps that include normalization and standardization within the data set aid to some extent in stabilizing learning dynamics, providing balance among different features during training. Improper scaling will result in inputs with divergent gradients hence unwieldy and erratic learning results. This stresses the necessity of data normalization as discussed in.

Relative entropy is another measure that is important for the scope for quantifying the differences between the model that arise as a consequence of using different random seeds. Unstable interpretations always have high relative entropy values, meaning that these interpretations must be worked on through the architecture of the model or the training strategies [9].

Another kind of prior that is useful is the Fourier-transform based attribution ones and these come in handy with genomic applications. These priors work by penalizing elements that are contributed by high frequency attribution scores hence improving the interpretability of the model while also making it strong enough to identify needed motifs without noise interference.

Additionally, tracking metrics like the Population Stability Index (PSI) provides valuable insights into shifts in data distributions over time. This helps identify when adjustments are needed to maintain model integrity in dynamic environments [8].

These mathematical techniques are essential not only for making a model more resilient to random fluctuations but also for ensuring consistent performance across various datasets and applications.

Month	AUC-ROC	F1-Score
April-2022	0.77	0.91
May-2022	0.78	0.91
June-2022	0.77	0.91
July-2022	0.76	0.9
August-2022	0.76	0.9
September-2022	0.75	0.89
October-2022	0.73	0.88
November-2022	0.73	0.87
December-2022	0.72	0.85

Fig 3: Model performance monitoring grid

5.2. Case Studies Illustrating Stability Issues

These case studies highlight the obstacles to achieving robustness within the confines of deep learning models, especially in the context of performing with a high dimensional feature space. One of the studies focused on the effect of alteration of a random seed on the model's performance. The findings were self-explanatory because even a single seed change could result in completely different outcomes. This often led to different explantations being offered for the same model predictions and features that were regarded as important, particularly in attention-based and gradient-based approaches [9]. The introduction of techniques such as Aggressive Stochastic Weight Averaging (ASWA) allowed to address some of these challenges, which boosted model stability across different seeds and achieved an average of 72 percent reduction of standard deviation of performance metrics.

In another study, the attention was devoted to the machine learning interatomic potentials (MLIPs) modeling in high dimensional spaces. It was determined that clustering based on error metrics can potentially mask dramatical instability for certain properties that are not easily predictable. This demonstrates the necessity of having specific approaches to control errors given the high dimensionality of the data for reliability improvement.

In addition, during concept drift, deep learning models for monitoring may be rendered inoperable due to the gradual decline of temporal resolution as models age and are subjected to severe shifts in distribution of data.

This makes retraining more difficult and can introduce new sources of instability due to factors like parameter selection and catastrophic forgetting.

These case studies highlight the complexity of stability issues in high-dimensional deep learning models and emphasize the need for developing robust techniques that can be applied across various fields.

6. Future Applications of Insights Gained from this Research

6.1. Implications for Computer Vision Technologies

Deep learning is thoroughly changing the face of many processes, improving both precision and performance metrics of multiple tasks. The introduction of Convolutional Neural Networks (CNNs) has enabled the extraction of intricate feature representations from high-dimensional visual data, improving image analysis and object recognition. This progress is evident on performance measures such as the Intersection over Union (IoU) for object identification and localization.

Likewise, there have been created diffusion models to fulfill certain particular requirements of traits of created visual content [6], and they are essential in modeling the data of high dimension. This is also known as the "shallow" far reaching phenomenon where extensive amounts of newly created data are able to be added to existing frameworks in place to utilize a wide array of data sources.

With self-supervised learning systems such as DINOv2 Self Supervised Learning Framework can perform unattended learning which is advantageous for automated systems with no labeled input data. This allows such systems to be more applicable to classification and segmentation of images with little tuning effort. This assists in the development of advanced self-supervised models for the construction of AI systems that automatically process information obtained from images.

Moreover, building specific models for Deep Learning (DL) based computer vision tasks is becoming easier, thanks to the integration of automated machine learning (AutoML) frameworks with Deep Learning. These frameworks eliminate the need for manual data labeling and hyperparameter tuning making it possible for non-experts to access and use advanced technology.

To sum up, one can conclude that the deep learning paradigm shift is fundamentally changing image analysis and fostering new opportunities for innovation in computer vision.

6.2. Relevance to Data Analysis Practices

Various industries benefit from the improved data analysis that deep learning models provide. These models are increasingly being used in various sectors, making their ability to effectively manage and understand high-dimensional data more apparent. These models are increasingly being adopted by businesses as they can learn through techniques such as federated learning, which allows models to adapt to different environments while still being accurate.

As many deep learning models, there exists the challenge of input data dimensionality. High dimensional spaces can result in simple problems becoming over complex, leading to overfitting, with greater computational costs. Thus, the need for dimensionality reduction techniques is necessary. Automatic encoders and Principal component analysis (PCA), for instance, provide ways to reduce data for retaining valuable aspects of it [2]. As a result, model performance gets enhanced further supporting the data generated by the models.

Deep learning algorithms, for example, have advanced agriculture in the area of plant disease diagnosis through image analysis. It shows how working with data can greatly improve disease detection and management for a specific field such as Agriculture. Lastly, consistent model performance evaluation is equally important to ensure that the model is effective over time, helping practitioners address potential issues before they become serious [5].

The economic aspects of deploying deep learning systems also need careful consideration. Understanding the costs involved in training and operating these models can help organizations optimize their resources and make better decisions [8]. Ultimately, integrating deep learning into data analysis not only boosts efficiency but also unlocks new opportunities for innovation across many fields.

7. Conclusions and Recommendations for Practitioners in the Field

7.1. Summary of Key Findings

Deep learning is often considered a black box due to its varying performance on different tasks, especially when working with a high dimensional feature space. To enable these models to work efficiently, there is a need to frequently evaluate them on a set of metrics, such as recall, precision, accuracy, and AUC-ROC scores. As Singh's case study showed, even highly successful models have a propensity to fail eventually, thus the need for consistent monitoring [7]. Cite PSI and CSI, which can help evaluate model stability and the existence of concept drift.

Some scholars believe that deep learning models can achieve performance benchmarks set by professional practitioners in certain domains. For example, one skeletal maturity assessment model's accuracy level was at par with that of veteran radiologists. However, there is significant work to be done regarding out-of-distribution, as these models require greater robustness to actual world scenarios.

To boost model performance, techniques like data scaling, feature selection, and regularization (To prevent overfitting) are commonly used, along with fine-tuning model architecture. Scaling the input features can significantly enhance both training efficiency and model effectiveness ([35]). Additionally, using diverse and representative training datasets is crucial, especially in areas like medical image analysis where the data needs to reflect the real-world diversity of cases.

Lastly, explainable AI tools such as XOmiVAE emphasize the importance of interpretability in deep learning. By making it easier to understand which features influence predictions, these tools help build trust among users and align model results with domain-specific knowledge [3].

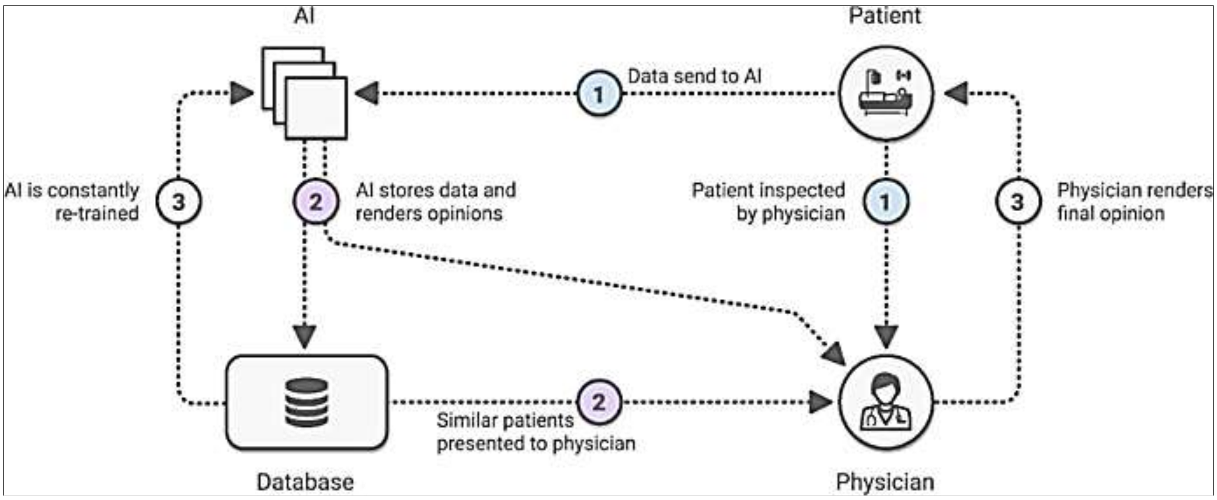

Fig 4: Clinical Deployment. An example workflow showing the positive compounding effect of AI-enhanced workflows, and the resultant trust that can be built. AI predictions provide immediate value to physicians, and improve over time as bigger datasets are collected.

7.2. Suggestions for Future Research Directions

Future research in the performance of deep learning models within high-dimensional environments should explore several key areas to improve efficiency, interpretability, and practical use. First, examining the limits of deep learning capabilities through comprehensive material synthesis could help identify when models reach a point of performance saturation. Understanding when adding more data no longer improves accuracy is crucial, as highlighted in.

Another major focus should be improving the interpretability of these models. Future studies should aim to clarify which features have the greatest impact on model predictions, particularly in fields like materials science and medical imaging. These considerations may facilitate improvement towards the design of experiments and the clinical reasoning processes.

Furthermore, increased attention should also be given to privacy-preserving techniques like federated learning. Federated learning allows data to be shared without actually exposing any sensitive information, leading to better model development with small datasets while being compliant to privacy laws.

Moreover, knowing how a model evolves concerning data changes—commonly known as concept drift—is critically important as well. Creating self-correcting techniques that allow models to remain accurate as the input data changes will be essential for sustained accuracy over time.

Lastly, the deployment costs of deep learning systems should be an important concern. As mentioned before, more work is needed to develop a framework that would enable the assessment of the balance between cost and efficiency in order to improve resource decisions (8). All these together will greatly improve the appreciation and application of deep learning models in complicated situations and settings.

8. Appendix: Mathematical Foundations Relevant to Deep Learning and Stability Analysis

As a rule of thumb, having a grasp of key mathematical concepts is crucial in deep learning since it helps with optimizing high-level model performance, especially in high feature spaces. One of them is embeddings which translates high-dimensional information to a lower dimension while retaining important aspects of the original data. Withnell *et al*. demonstrate the potential of autoencoders and variational autoencoders (VAEs) that compress the input data into more manageable latent variables, or features, for easier handling and analysis, as shown in XOmiVAE's (3) exploration.

Yet another example is dimensionality reduction. Some of the techniques include UMAP and t-distributed stochastic neighbor embedding (t-SNE), which reduce the data's dimensionality while maintaining the associated local data structures, which is useful in comprehension and interpretation of complex datasets (2). These techniques help with the so-called "curse of dimensionality," which refers to sparse data where traditional statistical methods do not work well.

A strong understanding of loss functions is equally important for training deep learning models effectively. The choice of loss function influences how a model learns from data, with common examples being mean squared error for regression tasks and cross-entropy loss for classification. Optimization algorithms like stochastic gradient descent (SGD) and Adam improve model performance by adjusting parameters during training.

Regularization techniques are also key for enhancing model stability and preventing overfitting in high-dimensional settings. By applying penalties to the model's magnitude or complexity through L1 and L2 regularization, models can generalize better to unseen data. These mathematical principles form the foundation of many deep learning methods.

References

1. Withnell E, Zhang X. XOmiVAE: an interpretable deep learning model for cancer classification using high-dimensional omics data. Jan 2021. Available from: https://academic.oup.com/bib/article/22/6/bbab315/6353242
2. Sindhwani N, Anand R, Meivel S, Shukla R, Yadav MP, Yadav V. Performance analysis of deep neural networks using computer vision. Oct 2021. https://eudl.eu/doi/10.4108/eai.13-10-2021.171318
3. Mashlakov A, Kuronen T, Lensu L, Kaarna A, Honkapuro S. Assessing the performance of deep learning models for multivariate probabilistic energy forecasting. Jan 2021. https://www.sciencedirect.com/science/article/pii/S0306261920317748
4. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, *et al*. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Mar 2021. Available from: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8
5. Tseng A, Shrikumar A, Kundaje A. Fourier-transform-based attribution priors improve the interpretability and stability of deep learning models for genomics. Feb 2020. Available from: https://proceedings.neurips.cc/paper/2020/hash/1487987e862c44b91a0296cf3866387e-Abstract.html
6. Gallagher B, Rever M, Loveland D, Mundhenk TN, Beauchamp B, Robertson E, *et al*. Predicting compressive strength of consolidated molecular solids using computer vision and deep learning. Jan 2020. Available from: https://www.sciencedirect.com/science/article/pii/S0264127520300745
7. Ahmad A, Saraswat D, Aggarwal V, Etienne A, Hancock B. Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. Jan 2021. Available from: https://www.sciencedirect.com/science/article/pii/S0168169921000995
8. Chen K, Chen H, Zhou C, Huang Y, Qi X, Shen R, *et al*. Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. Mar 2020. Available from: https://www.sciencedirect.com/science/article/pii/S004313541931231X
9. Waheed H, Hassan S, Aljohani NR, Hardman J, Alelyani S, Nawaz R. Predicting academic performance of students from VLE big data using deep learning models. Jan 2020. Available from: https://www.sciencedirect.com/science/article/pii/S0747563219304017.