Towards Robust Deep Learning

Deep Latent Variable Modeling against Out-of-Distribution and Adversarial Inputs

Glazunov, M.

DOI
10.4233/uuid:beacb059-da33-42ed-a094-8d7674f2ec26

Publication date
2025

Document Version
Final published version

Citation (APA)
Glazunov, M. (2025). *Towards Robust Deep Learning: Deep Latent Variable Modeling against Out-of-Distribution and Adversarial Inputs*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:beacb059-da33-42ed-a094-8d7674f2ec26

Propositions

accompanying the dissertation

Towards Robust Deep Learning

Deep Latent Variable Modeling against Out-of-Distribution and Adversarial Inputs

by

Mikhail GLAZUNOV

1. Control over the compactness of the latent space allows for a balance between the model's expressivity and robustness. *(This proposition pertains to this dissertation).*
2. Out-of-distribution inputs can serve as effective availability attacks on deep neural networks, similar to adversarial examples. *(This proposition pertains to this dissertation).*
3. Holes in the latent space of variational autoencoders, often considered undesirable in literature, enable the model to deal with both out-of-distribution and adversarial inputs. *(This proposition pertains to this dissertation).*
4. Prediction overconfidence of deep neural classifiers directly follows from the limiting nature of any classification task. *(This proposition pertains to this dissertation).*
5. Based on the foundational principles of the superposition theorem, Kolmogorov-Arnold Networks (KANs) will eventually replace the Deep Neural Networks (DNNs) that are extensively utilized at present.
6. Explainable AI will never explain deep neural networks.
7. To keep pace with the rapid advancements in scientific knowledge, it is crucial to employ AI systems that extract, summarize, and integrate new information, thereby ensuring we effectively harness the expansive array of available information.
8. Discouraging the publication of negative results in top-ranking venues hinders scientific progress.
9. The introduction of an index, similar to the Hirsch index, recognizing the contributions of researchers as peer reviewers, improves the quality and speediness of the paper review process.
10. Naps for adults at work are more effective at increasing productivity than using AI tools.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotor prof. dr. ir. R. L. Lagendijk and the co-promotor dr. D.M.J. Tax.