

# Widespread machine learning methods behind 'link prediction' are performing very poorly, researchers find

February 12 2024

---



Credit: CC0 Public Domain

As you scroll through any social media feed, you are likely to be prompted to follow or friend another person, expanding your personal

network and contributing to the growth of the app itself. The person suggested to you is a result of link prediction: a widespread machine learning (ML) task that evaluates the links in a network—your friends and everyone else's—and tries to predict what the next links will be.

Beyond being the engine that drives social media expansion, link prediction is also used in a wide range of scientific research, such as predicting the interaction between genes and proteins, and is used by researchers as a benchmark for testing the performance of new ML algorithms.

New research from UC Santa Cruz Professor of Computer Science and Engineering C. "Sesh" Seshadhri [published](#) in the journal *Proceedings of the National Academy of Sciences* establishes that the metric used to measure link prediction performance is missing crucial information, and link prediction tasks are performing significantly worse than popular literature indicates.

Seshadhri and his co-author Nicolas Menand, who is a former UCSC undergraduate and masters student and a current Ph.D. candidate at the University of Pennsylvania, recommend that ML researchers stop using the standard practice metric for measuring link prediction, known as AUC, and introduce a new, more comprehensive metric for this problem. The research has implications for trustworthiness around decision-making in ML.

## **AUC's ineffectiveness**

Seshadhri, who works in the fields of theoretical computer science and [data mining](#) and is currently an Amazon scholar, has done previous research on ML algorithms for networks. In this previous work, he found certain mathematical limitations that were negatively impacting [algorithm](#) performance, and in an effort to better understand the

mathematical limitations in context, dove deeper into link prediction due to its importance as a testbed problem for ML algorithms.

"The reason why we got interested is because link prediction is one of these really important scientific tasks which is used to benchmark a lot of machine learning algorithms," Seshadhri said.

"What we were seeing was that the performance seemed to be really good... but we had an inkling that there seemed to be something off with this measurement. It feels like if you measured things in a different way, maybe you wouldn't see such great results."

Link prediction is based on the ML algorithm's ability to carry out low dimensional vector embeddings, the process by which the algorithm represents the people within a network as a mathematical vector in space. All of the machine learning occurs as mathematical manipulations to those vectors.

AUC, which stands for "area under curve" and is the most common metric for measuring link prediction, gives ML algorithms a score from zero to one based on the algorithm's performance.

In their research, the authors discovered that there are fundamental mathematical limitations to using low dimensional embeddings for link predictions, and that AUC can not measure these limitations. The inability to measure these limitations caused the authors to conclude that AUC does not accurately measure link prediction performance.

Seshadhri said these results call into question the widespread use of low dimensional vector embeddings in the ML field, considering the mathematical limitations that his research has surfaced on their performance.

## Leading methods fall short

The discovery of AUC's shortcomings led the researchers to create a new metric to better capture the limitations, which they call VCMPR. They used VCMPR to measure 12 ML algorithms chosen to be representative of the field, including algorithms such as DeepWalk, Node2vec, NetMF, GraphSage, and graph benchmark leader HOP-Rec, and found that the link prediction performance was worse using VCMPR as the metric rather than AUC.

"When we look at the VCMPR scores, we see that the scores of most of the leading methods out there are really poor," Seshadhri said. "It looks like they're actually not doing a good job when you measure things a different way."

The results also showed that not only was performance lower across the board, but some of the algorithms that performed worse than other algorithms when measured with AUC in turn performed better than the cohort with VCMPR, and vice versa.

## Trustworthiness in machine learning

Seshadhri suggests that ML researchers use VCMPR to benchmark the link prediction performance of their algorithms, or at the very least stop using AUC as their measure. As metrics are so tightly connected to decision-making in ML, using a flawed system to measure performance could lead to flawed decision-making about which algorithms to employ in real-world ML applications.

"Metrics are so closely tied to what we decide to deploy in the real world—people need to have some trust in that. If you have the wrong way of measuring, how can you trust the results?" Seshadhri said. "This

paper is in some sense cautionary: we have to be more careful about how we do our machine learning experiments, and we need to come up with a richer set of measures."

In academia, using an accurate metric is crucial to creating progress in the ML field.

"This is in some sense a bit of a conundrum for scientific progress. A new result has to supposedly be better than everything previously, otherwise it's not doing anything new—but that all depends on how you measure it."

Beyond [machine learning](#), there are researchers across a wide range of fields who use link prediction and ML to conduct their research, often with profound potential impact. For example, some biologists use link prediction to determine which proteins are likely to interact as a part of drug discovery. These biologists and other researchers outside of ML depend on ML experts to create trustworthy tools, as they often cannot become ML experts themselves.

While he thinks these results may not be a huge surprise to those deeply involved in the field, he hopes that the larger community of ML researchers, and particularly graduate and Ph.D. students who use the current literature to learn [best practices](#) and common wisdom about the field, will take note of these results and take caution in their work. He sees this research that presents a skeptical view to be in somewhat contrast to a dominant philosophy in ML, which tends to accept a set of metrics and focuses on "pushing the bar" when it comes to progress in the field.

"It's important that we have the skeptical view, are trying to understand deeper, and are constantly asking ourselves 'Are we measuring things correctly?'"

**More information:** Menand, Nicolas et al, Link prediction using low-dimensional node embeddings: the measurement problem, *Proceedings of the National Academy of Sciences* (2024). [DOI: 10.1073/pnas.2312527121](https://doi.org/10.1073/pnas.2312527121). [doi.org/10.1073/pnas.2312527121](https://doi.org/10.1073/pnas.2312527121)

Provided by University of California - Santa Cruz

Citation: Widespread machine learning methods behind 'link prediction' are performing very poorly, researchers find (2024, February 12) retrieved 27 April 2024 from <https://techxplore.com/news/2024-02-widespread-machine-methods-link-poorly.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.