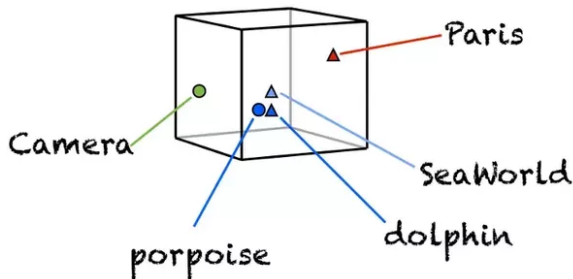


Методы ИИ в естественнонаучных областях.
«Слабый» и «сильный» ИИ. Проблема
интерпретируемости современных систем ИИ

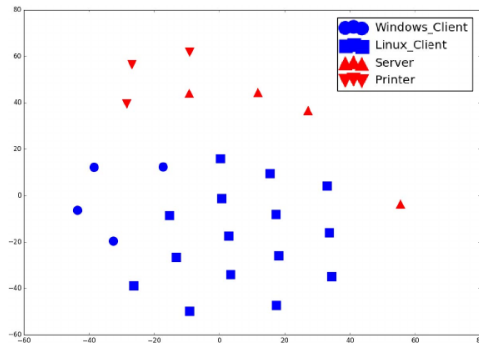
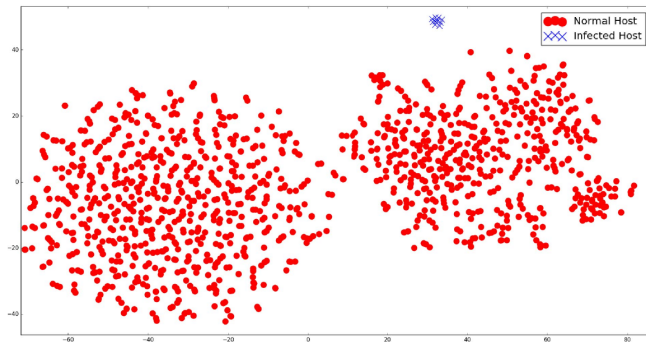
Зубюк Андрей Владимирович
zubjuk@physics.msu.ru

<http://NeuroFuzzy.Phys.MSU.ru>



- ▶ Каждому слову ставится в соответствие вектор: Word $\xrightarrow{\text{to} - 2}$ Vec.
- ▶ Векторы, соответствующие словам, подбираются так, чтобы словам, имеющим близкий смысл, соответствовали близкие векторы.
- ▶ Показателем близости смысла считается употребление в схожих контекстах. Контекст слова понимается тривиальным образом — как несколько слов, окружающих выбранное (например, 5 предшествующих слов и 5 последующих).
- ▶ Подбор векторов осуществляется методами машинного обучения на основе корпуса текстов.

M. Ring et al., "Ip2vec: Learning similarities between ip addresses", in **2017 IEEE International Conference on Data Mining Workshops (ICDMW)**, 2017, pp. 657–666. DOI: [10.1109/ICDMW.2017.93](https://doi.org/10.1109/ICDMW.2017.93)



K. K. Yang et al., “Learned protein embeddings for machine learning”, **Bioinformatics** (Oxford, England), pp. 2642–2648, 2018. DOI: [10.1093/bioinformatics/bty178](https://doi.org/10.1093/bioinformatics/bty178)

Исходный код: https://github.com/fhalab/embeddings_reproduction/

Machine-learning models trained on protein sequences and their measured functions can **infer biological properties of unseen sequences without requiring an understanding of the underlying physical or biological mechanisms**. Such models enable the prediction and discovery of sequences with optimal properties. Machine-learning models generally require that their inputs be vectors, and **the conversion from a protein sequence to a vector representation** affects the model’s ability to learn. We propose to learn embedded representations of protein sequences that take advantage of the vast quantity of unmeasured protein sequence data available. These embeddings are low-dimensional and can greatly simplify downstream modeling.

C. Réda, E. Kaufmann, and A. Delahaye-Duriez, “Machine learning applications in drug development”, **Computational and Structural Biotechnology Journal**, vol. 18, pp. 241–252, 2020. DOI: [10.1016/j.csbj.2019.12.006](https://doi.org/10.1016/j.csbj.2019.12.006)

124 ссылки.

Highlights:

- ▶ Applications of sequential learning and recommender systems to pharmaceuticals.
- ▶ Review of Machine Learning methods in drug discovery, testing and repurposing.
- ▶ Survey of available genomic data and feature selection methods for drug development.

G. E. Karniadakis **et al.**, “Physics-informed machine learning”, **Nature Reviews Physics**, vol. 3, pp. 422–440, 2021

Обучение алгоритмов решения обратных задач в физике, в т. ч. задач анализа данных физического эксперимента.

Z. Ren **et al.**, “Physics-guided characterization and optimization of solar cells using surrogate machine learning model”, in **2019 IEEE 46th Photovoltaic Specialists Conference (PVSC)**, 2019, pp. 3054–3058. DOI: [10.1109/PVSC40753.2019.8980715](https://doi.org/10.1109/PVSC40753.2019.8980715)

Characterization, material parameter extraction and subsequent optimization of solar cell devices **is a highly time-consuming and complex procedure**. In this work, we propose **a method for quick extraction of limiting material parameters** in solar cell devices using a surrogate, physics-embedded, neural network model. This surrogate model, implemented by an autoencoder architecture trained with a physical numerical model, allows to quickly extract the device parameters of interest at a certain process condition

Проект паспорта новой научной специальности «Искусственный интеллект и машинное обучение»:

https://drive.google.com/drive/folders/1xqoWINSPPH48_IA2Iw1uuWt3qkMQc5E0:

Исследования в области «сильного ИИ», включая формирование понятийной базы и элементов математического формализма, необходимых для построения алгоритмического аппарата.

Artificial general intelligence (AGI) is the hypothetical ability of an intelligent agent to understand or learn any intellectual task that a human being can. It is a primary goal of some artificial intelligence research and a common topic in science fiction and futures studies. AGI can also be referred to as strong AI, full AI, or general intelligent action.

EXplainable **A**rtificial **I**ntelligence (XAI) is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms. Explainable AI is used to describe an AI model, its expected impact and potential biases. It helps characterize model accuracy, fairness, transparency and outcomes in AI-powered decision making. Explainable AI is crucial for an organization in building trust and confidence when putting AI models into production. AI explainability also helps an organization adopt a responsible approach to AI development.

Одно из направлений развития XAI — совмещение (в некотором смысле) «классических» методов ИИ с плохо интерпретируемыми современными (в т. ч. нейросетевыми).

Классические	Нейросетевые
Интерпретируемость	Качество

При прохождении теста укажите e-mail и фамилию, имя, отчество, которые вы указали при регистрации на курс. По этим данным будут суммироваться результаты всех ваших тестов в семестре. После ответа на тест вам на почту должно прийти письмо с вашими ответами.

https://docs.google.com/forms/d/e/1FAIpQLSdFBQfNwxVk5-2NBKQiMv4ild_by40iRlE51Gm_btNlHr9fSQ/viewform

https://neurofuzzy.phys.msu.ru/redmine/projects/ai_in_natural_science/news

