

**А. К. Горшенин** (Москва, ФИЦ ИУ РАН). **Вероятностно-статистические методы и искусственный интеллект: о современном состоянии и тенденциях развития.**

УДК 004.8+519.25

DOI [https://doi.org/10.52513/08698325.2022\\_29\\_3\\_1](https://doi.org/10.52513/08698325.2022_29_3_1)

*Резюме:* В докладе рассмотрена тесная взаимосвязь методов вероятности, статистики и алгоритмов машинного обучения и искусственного интеллекта. Приведен обзор как хорошо известных результатов в этой области, так и наиболее современных работ, в том числе касающихся вопросов расширения признакового пространства при обучении, а также вероятностных трансформеров.

*Ключевые слова:* Статистическое обучение, искусственный интеллект, машинное обучение, глубокие нейронные сети, вероятностные модели, расширение признакового пространства.

Методы машинного обучения и алгоритмы искусственного интеллекта в последнее десятилетие позволили добиться впечатляющих результатов в различных научных областях, успешно справляясь с задачами моделирования и анализа реальных данных. В частности, достигнуты успехи и при отсутствии детерминированности в поведении реальных процессов из-за стохастического характера интенсивностей потоков событий, неоднородности или нестационарности изучаемых закономерностей, неполноты получаемой информации. Все это условия, традиционно ведущие к необходимости построения вероятностно-статистических моделей.

В докладе рассматривается тесная взаимосвязь методов искусственного интеллекта с вероятностными подходами [1], начиная с классических результатов теории распознавания образов [2], основ статистического обучения В.Н.Валника и А.Я.Червоненкиса [3], К.Хорника [4]. Обсуждаются современные применения методов вероятности и статистики в задачах обучения с подкреплением [5] и в развитии известного EM-алгоритма — метода, получившего широкое распространение для построения оценок максимального правдоподобия параметров распределений, который однако, как было недавно продемонстрировано Б.Коско [6] обобщает классический подход к обучению нейронных сетей — метод обратного распространения ошибки [7]. Впоследствии был создан и нейронный EM-алгоритм [8], а также архитектура, названная глубокими гауссовскими смесями [9]. Рассмотрены примеры вероятностных подходов и в популярных современных нейросетевых архитектурах — трансформерах [10].

Приведены результаты исследований, проводимых в ФИЦ ИУ РАН, в области математического моделирования реальных данных с помощью стохастических дифференциальных уравнений совместно с использованием различных классов нейронных сетей. Показано, что использование характеристик вероятностно-статистических моделей в качестве дополнительных обучающих признаков позволяет повысить точность нейросетевого прогнозирования в среднем на 16% для краткосрочных и на 12% для среднесрочных прогнозов по сравнению с известными методами в целом ряде прикладных направлений, включая физику плазмы и климатологические приложения [11].

Продемонстрировано [12], что добавление дополнительных признаков — информации о вероятностном распределении приращений данных — в ансамблевой LSTM-

модели позволяет превосходить как базовые методов алгоритмической торговли по финансовым индикаторам (преимущество до 32.2% на тестовых данных), так и известные подходы на основе нейронных сетей без статистического расширения признакового пространства (преимущество до 23.3%).

Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика» ФИЦ ИУ РАН, Москва).

Работа выполнена при финансовой поддержке Российского научного фонда, проект 22-11-00212.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Bzdok D., Altman N., Krzywinski M.* Statistics versus machine learning. — *Nature Methods*, 2018, v. 15, is. 4, p. 233–234.
2. *Fukunaga K.* Introduction to statistical pattern recognition. New York: Academic Press, 1990, 594 p.
3. *Vapnik V.* Statistical Learning Theory. New York: Wiley, 1998, 768 p.
4. *Hornik K., Stinchcombe M., White H.* Multilayer feedforward networks are universal approximators. — *Neural Networks*, 1989, v. 2, is. 5, p. 359–366.
5. *Sutton R.S., Barto A.G.* Reinforcement Learning: An Introduction. Cambridge: The MIT Press, 2018, 344 p.
6. *Audhkhasi K., Osoba O., Kosko B.* Noise-enhanced convolutional neural networks. — *Neural Networks*, 2016, v. 78, p. 15–23.
7. *Rumelhart D.E., Hinton G.E., Williams R.J.* Learning Internal Representations by Error Propagation. — *Parallel Distributed Processing*, 1986, v. 1, p. 318–362.
8. *Greff K., van Steenkiste S., Schmidhuber J.* Learning Internal Representations by Error Propagation. — *Advances in Neural Information Processing Systems (NIPS'17)*, 2017, p. 6694–6704.
9. *Viroli C., McLachlan G.J.* Deep Gaussian Mixture Models. — *Statistics and Computing*, 2019, v. 29, p. 43–51.
10. *Tang B., Matteson D.S.* Probabilistic Transformer for Time Series Analysis. — *Advances in Neural Information Processing Systems (NeurIPS 2021)*, 2021, p. 1–17.
11. *Gorshenin A.K., Kuzmin V.Yu.* Statistical Feature Construction for Forecasting Accuracy Increase and its Applications in Neural Network Based Analysis. — *Mathematics*, 2022, v. 10, is. 4, art. № 589.
12. *Gorshenin A.K., Vilyaev A.L.* Finite Normal Mixture Models for the Ensemble Learning of Recurrent Neural Networks with Applications to Currency Pairs. — *Pattern Recognition and Image Analysis*, 2022, v. 32, № 4, p. 780–792.

UDC 004.8+519.25

DOI [https://doi.org/10.52513/08698325\\_2022\\_29\\_3\\_1](https://doi.org/10.52513/08698325_2022_29_3_1)

**Gorshenin A. K.** (Moscow, Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, Lomonosov Moscow State University). **Probabilistic-statistical methods and artificial intelligence: on the current state and trends.**

*Abstract:* The report discusses the close relationship between methods of probability, statistics as well as machine learning and artificial intelligence algorithms. A review of both well-known results in this area and the most recent papers, including those related to the issues of the feature expansion as well as probabilistic transformers, is given.

*Keywords:* Statistical learning, artificial intelligence, machine learning, deep neural networks, probabilistic models, feature expansion