

УДК 004.85

DOI

**Уткин Л.В.,**

доктор технических наук, профессор,  
vlad2tu@yandex.ru<sup>1</sup>

**Заборовский В.С.,**

доктор технических наук, профессор,  
vlad2tu@yandex.ru<sup>1</sup>

*Санкт-Петербургский политехнический университет  
Петра Великого  
195251, Санкт-Петербург, ул. Политехническая, 29<sup>1</sup>*

## **О ПРОБЛЕМЕ МАШИННОГО ОБУЧЕНИЯ ГИБРИДНЫХ СУПЕРКОМПЬЮТЕРОВ КАК ЦЕНТРОВ КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ**

***Аннотация.** Бурное развитие вычислительных технологий и создание универсальных генеративных моделей-трансформеров ставит перед исследователями фундаментальный вопрос: при каких условиях и для каких задач можно ли заменить прикладные программы вычислений, написанные с использованием формальных языков, на промпт-запросы к пред-обученным мультимодальным большим языковым моделям. В статье, опираясь на абдуктивный вывод о том, что «физическое» порождает «информационное», а «информационное вызывает вычислительное», рассматривается механизм влияния процессов машинного обучения на производительность гибридных суперкомпьютерных кластеров, работающих в режиме центров коллективного пользования, подтверждая актуальность тезиса Ж. Л. Лагранжа – «не будем спорить, давайте посчитаем».*

***Ключевые слова:** суперкомпьютеры, машинное обучение, модель выживаемости, экзо-интеллект, объяснительный интеллект, доказательство, вероятность.*

*Время похоже на дорогу: она не исчезает после того,  
как мы прошли по ней, и не возникает сию секунду,  
открываясь за поворотом  
Робнерт Бартини*

### **1. Введение**

Развитие компьютерных наук и создание глобальной информационной сети Интернет радикально изменили методологию проведения научных исследований, позволив ученым не только индивидуально изучать различные физические процессы методами математического моделирования, но описывать реальность, используя так называемые «большие» данные, в частности, представленные в формате мультимодальных корпусов текстов. Такие данные являются основой для создания трансформерных моделей-универсальных интерфейсов для обработки сложных запросов и генерации на них максимально правдоподобных со статистической точки зрения ответов. Языковые трансформеры, построенные на статистическом непараметрическом механизме внимания, не могут обучаться непосредственно при взаимодействии с конкретным пользователем, учитывая при этом текущий контекст запроса и темпоральные ограничения. Поэтому их применение без дополнительной системы объяснения сгенерированных результатов несет опасность подмены описания объектов реальности на основе законов сохранения - первых принципов физики, описанием на основе корреляций и индуктивных правил. В этих условиях особую актуальность для эффективной работы суперкомпьютерных центров коллективного пользования приобретают фундаментальные вопросы алгоритмической вычислимости (разрешимости) отдельных особых состояний физических процессов и объектов, к которым, в том числе, относятся и сами современные компьютеры.

Почти сто лет А. Тьюрингом (Alan Turing) была сформулирована и доказана теорема об алгоритмической неразрешимости проблемы вычисления времени «останова» (Halting problem) компьютерной программы, которая управляет процессами изменения состояния конечного автомата - «машины Тьюринга» - универсальная модель вычислений, которая может имитировать любое другое вычислительное устройство. Проблема «останова» формулируется просто: требуется составить программу  $U$ , которая бы по любому подаваемому ей на вход файлу  $X$ , содержащему текст программы на каком-нибудь языке программирования, определяла бы, остановится ли когда-нибудь программа из файла  $X$  в процессе своей работы, получив на вход известные данные, или "зациклится".

В дальнейшем были найдены доказательства алгоритмической неразрешимости различных массовых проблем и сформулирован тезис Чёрча - Тьюринга, который утверждает, что любые вычислительные процессы, реализуемые на современном компьютере или любой другой универсальной вычислительной машине, могут быть смоделированы с помощью машины Тьюринга. Другими словами, была высказана абдуктивное положение о том, что осуществимые на практике вычисления могут быть представлены в формальных дедуктивных

системах, выразительные возможности которых ограничены теоремами Геделя. Но именно основе этого положения стала развиваться теория вычислительной сложности, основанная на классификации компьютерных проблем с использованием дескрипторов, характеризующего сложность вычислений.

**Проблемы интеллектуализации вычислений.** Стремительная коэволюция технологий квантовой электроники, теории категорий и машинного обучения на основе механизма внимания [3], наделяет современные компьютеры способности к адаптации и самоорганизации, что позволяет пересмотреть классификацию проблем сложности вычислений, которая изначально опиралась на формализм машины Тьюринга. Использование генеративных механизмов, эмитирующих феномен аутопоэзиса [1,2] вычисленных данных и моделируемых явлений-понятий позволяет (Рис.1) позволяет уточнить формулировку проблемы самоприменимости вычислительных процессов, которые при использовании формализма Тьюринга в общем случае алгоритмически неразрешимы.

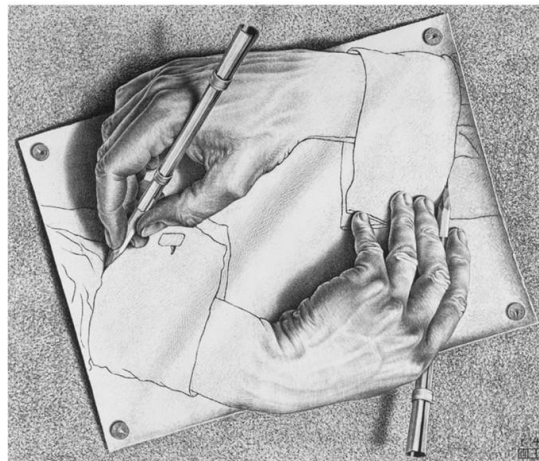


Рисунок 1. «Рисующие руки» — литография Эшера, 1948 г.

Пополнение машины Тьюринга механизмами машинного обучения позволяет ввести новый класс сложности, а именно класс ML который соответствует алгоритмам P класса, полученным из NP алгоритмов, дополненных средствами искусственного интеллекта AI (artificialintelligent). Такое уточнение можно представить в виде мнемонической формулы

$$NP+AI \Rightarrow ML = P$$

В результате высокопроизводительные цифровые компьютеры, дополненные средствами машинного обучения на базе генеративных моделей, с точки зрения законов физики можно рассматривать как электронные устройства-трансформеры, участвующие в преобразовании свободной электрической энергии в

процессы передачи информации, содержащейся в программном коде, в данные, которые кодируют различные состояния арифметико-логических устройств компьютера. При этом встраивание в компьютеры механизма объяснения может стать поворотной точкой (не Turing, а turningpoint) в трансформации компьютерных наук, как сферы реализации необратимых физических процессов, развивающихся в соответствии с формулой Ландауэра [3], в новую область модельно обратимых процессов обработки информации (inFORMAtion), форма которой кодирует ее амодельное или трансцендентное содержание (Рис. 2).

В основе новых технологий лежит принцип «обучения машин», который рассматривается как процесса решения обратных задач класса «воспринимаемые данные – алгоритм вычисления», сохранения в памяти суперкомпьютера информацию о результатах вычислений, которые могут автоматически размечаются диспетчером суперкомпьютерных ресурсов и в дальнейшем рассматриваться как одна из форм представления обучающей выборки системы управления процессами вычислений. Пользователь такой систем приобретает способность получать результаты вычислений не только на основе информации, носителем которой является его прикладная программа, но также и на основе результатов предсказательного моделирования потенциально возможных последствий, которые проявятся в «будущем», но в текущий момент могут быть вычислены на основе результатов полученных «в прошлом». Такая информационная форма ретро причинности [6, 7] или виртуальная временная инверсия причинно-следственных отношений, основанная на результатах предсказательном моделировании, является конструктивным вариантом «дооснащения» интеллектуальных процессов механизмами редукции, восходящих к работам Скулема [10]. Для того, чтобы объекты или процессы, были включены в процессы предсказательного моделирования, они должны быть предварительно «подменены» соответствующим понятиями, модальным символом или амодальной формой репрезентации. Такая модель процессов вычислений, дополненная средствами машинного обучения, допускает естественное описание результатов вычислений в терминах теории категорий [14] и различных мыслимых формах.

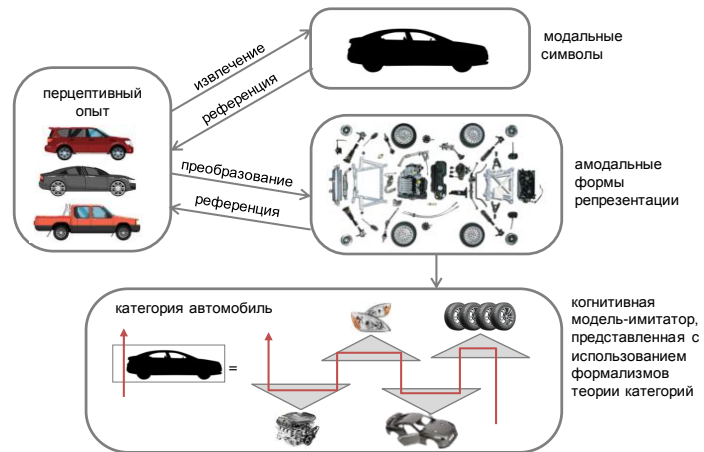


Рисунок 2. Модели объяснительной рациональности на основе формализма теории категорий перцептивных, модальных и амодальных репрезентаций

Отличие процесса «интеллектуальных» вычислений со встроенными средствами машинного обучения в том, что фактор-множество результатов вычислений может не образовывать алгебраическую системы, поэтому вычисление результатов с точностью, характерной для числовых функций, не всегда возможно (Рис.3)



Восприятие через: а) слова-понятия б) функции-числа

Рисунок 3. Восприятие реальности:  
а) фактор-множества понятий, б) функции времени  $u(t)$

Заметим, что именно слова-понятия, наделенные семантическим содержанием, формируют целостное представление о воспринимаемых явлениях, сохраняя их как темпоральную, так и функциональную упорядоченность.

**3. Машинное обучение диспетчера суперкомпьютера.** Все предыдущие рассуждения могут быть применены к таким сложным системам как современный гибридный суперкомпьютерный кластер, работающий в режиме центра коллективного пользования. Такой кластер-это техническая система, которая одновременно выполняет квадриллионы (число с 15 нулями) машинных операций, используемых в алгоритмах решения сложных математических задачи обработки

огромных объемов данных, генерируемых в процессе решения прикладных и системных задач. Пользователи суперкомпьютеров, число которых неуклонно растет с каждым годом, заинтересованы в получении результатов своих вычислений с максимально возможной производительностью и точностью. По мере увеличения производительности суперкомпьютеров, особенно при использовании гибридных архитектур, возрастает и сложность их использования.

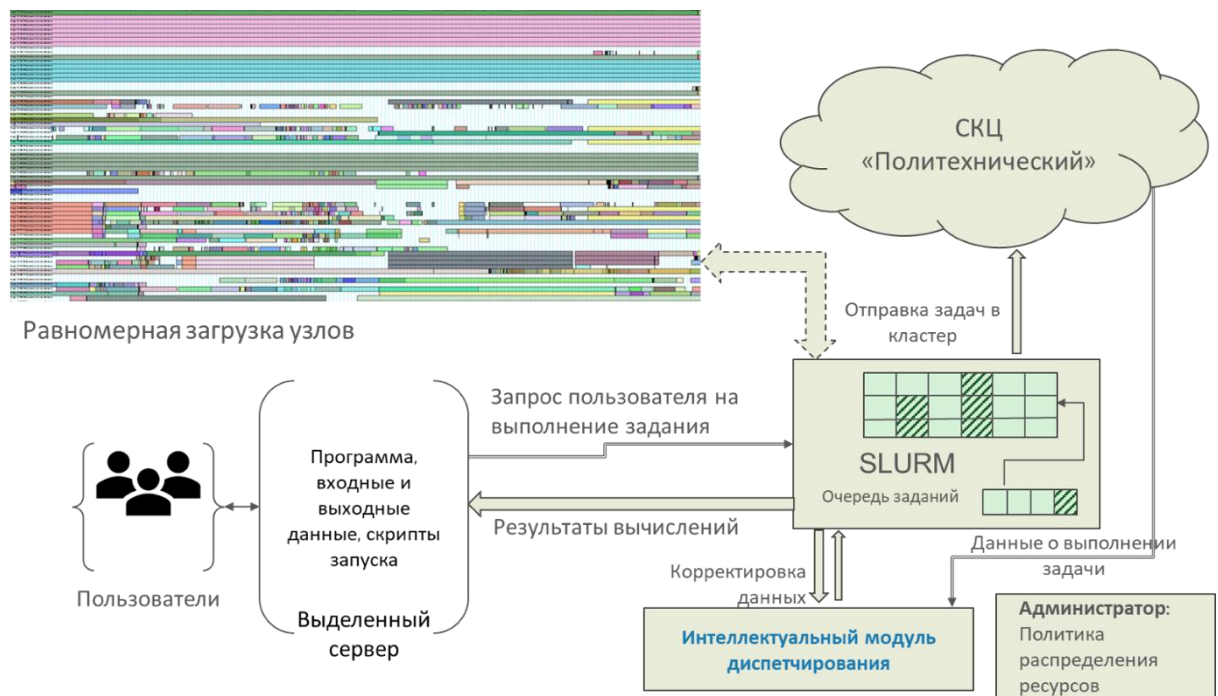


Рисунок 4. Характер загрузки узлов гибридного суперкомпьютерного кластера, оснащенного модулем машинного обучения диспетчера ресурсов

Для достижения наибольшей производительности суперкомпьютера, которую можно оценить как относительную долю успешно решенных прикладных задач за выделенный интервал времени, важную роль играет диспетчер – планировщик процессов вычислений, распределяющий имеющиеся вычислительные ресурсы между прикладными задачами с учетом требований к точности решений, объему используемой памяти и числу процессорных ядер. В большинстве современных суперкомпьютеров эти параметры задаются сам пользователем, который при этом опирается на свой весьма ограниченный опыт и объем знаний по программированию гибридных суперкомпьютеров. Поэтому именно система планирования ресурсов суперкомпьютера является той его частью, функциональность которой может быть эффективно дополнена средствами машинного обучения.

Сведение задачи управления производительностью суперкомпьютера к анализу модели выживаемости прикладных задач открывает широкие

возможности для статистической характеристики процессов распределения ресурсов в контексте применения методов машинного обучения системы планирования и распределения вычислительных ресурсов [7]. Следует отметить, важной особенностью моделей выживаемости, которые могут использоваться для решения задачи прогнозирования не только точечных оценок данных, как это происходит в большинстве моделей машинного обучения, но для оценки вероятностей получения этих оценок. На Рис. 4 показана временное разнообразие времени исполнения различных задач в СКЦ «Политехнический», которое

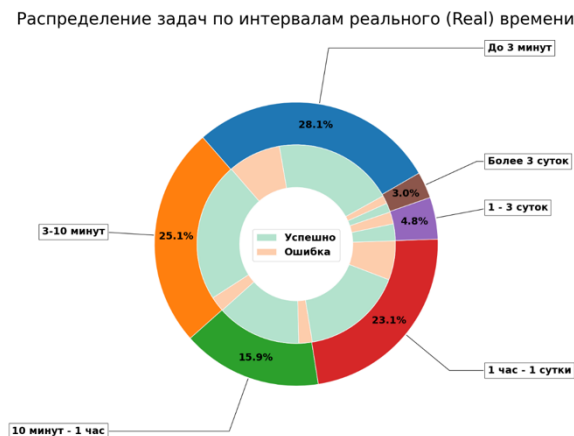


Рисунок 5. Распределение задач по интервалам времени из выполнения.

позволяет оценить долю успешно завершённых задач, так как на Рис.4 каждый сектор круговой диаграммы разбит на два подсектора, а которых указаны дескрипторы «успешно» и «ошибка». Учитывая, что СКЦ «Политехнический» функционирует в режиме центра коллективного пользования, для реализации процессов машинного обучения системы диспетчерского управления необходимо учитывать информации о времени выполнения прикладных задач, относящихся к различным областям знаний.

Результаты анализа статистического значимости различий функций распределения времени, необходимого для успешного выполнения прикладных задач в различных областях знаний, могут быть построены на оценке достоверности попарных различий изучаемых распределений в выделенных областях знаний. Именно установленный факт различия указывает на то, какие признаки прикладных задач можно выбрать в качестве фактора группировки экспериментальных данных для построения на них основе обучающих выборок для системы машинного обучения диспетчера вычислительных ресурсов. Такой выбор позволяет учесть особенностей вычислительных процессов в гибридной суперкомпьютерной среде с точки зрения успешного выполнения прикладных задач, что можно сделать с помощью вычислимой функции выживания прикладной задачи, используемой в системе машинного

обучения диспетчера ресурсов. (Рис.5). Для формализации предложенного подхода к оценке производительности, рассмотрим обучающее множество  $D$ , состоящее из  $n$  троек  $(\mathbf{x}_i, T_i, \delta_i)$ ,  $i = 1, \dots, n$ , где каждая тройка характеризует объект,  $\mathbf{x}_i \in \mathbf{R}^m$  - вектор атрибутов или признаков;  $T_i$  - время до события  $i$ -го объекта,  $\delta_i$  - индикатор события,  $\delta_i = 1$ , если событие наблюдалось (нецензурированное наблюдение),  $\delta_i = 0$ , если событие не наблюдалось (цензурированное наблюдение). Цель - оценить время до события  $T$  на основе  $D$  для нового объекта, имеющего вектор атрибутов  $\mathbf{x}$ .

Ключевыми понятиями в анализе выживаемости являются функции выживаемости и риска. Функция выживаемости  $S(t|\mathbf{x})$  есть вероятность того, что событие, связанное с объектом  $\mathbf{x}$ , не произошло до момента времени  $t$ . Важным понятием является кумулятивная функция риска  $H(t|\mathbf{x})$ , которая выражается через функцию выживаемости по формуле  $H(t|\mathbf{x}) = -\ln(S(t|\mathbf{x}))$ .

Одной из моделей выживаемости является модель Кокса, для которой кумулятивная функция риска определяется как

$$H(t|\mathbf{x}, \mathbf{b}) = H_0(t) \cdot \exp(\mathbf{b}^T \mathbf{x}),$$

где  $H_0(t)$  - базовая кумулятивная функция риска, определяемая, например, как оценка Каплана-Мейера или Нельсона-Аалена,  $\mathbf{b}$  - вектор параметров модели.

Главная особенность модели Кокса заключается в том, что функция риска или функция выживаемости зависят от линейной комбинации атрибутов, что позволит в дальнейшем использовать эту модель для объяснения предсказаний моделей выживаемости. Однако модель Кокса не учитывает взаимных расстояний между точками в обучающей выборке. Эту проблему решает оценка Берана [6], в соответствии с которой функция выживаемости определяется следующим образом:

$$S_B(t|\mathbf{x}) = \prod_{T_i \leq t} \left\{ 1 - \frac{W(\mathbf{x}, \mathbf{x}_i)}{1 - \sum_{j=1}^{i-1} W(\mathbf{x}, \mathbf{x}_j)} \right\}^{\delta_i}.$$

Здесь вес  $W(\mathbf{x}, \mathbf{x}_i)$  характеризует близость векторов  $\mathbf{x}, \mathbf{x}_i$ . Например, вес может определяться через операцию softmax:  $W(\mathbf{x}, \mathbf{x}_i) = \text{softmax}(\|\mathbf{x} - \mathbf{x}_i\|^2 / \tau)$ , где  $\tau$  - параметр оценки. Интересно отметить, что модель Каплана-Мейера является частным случаем оценки Берана, когда все веса одинаковы и равны  $1/n$ .

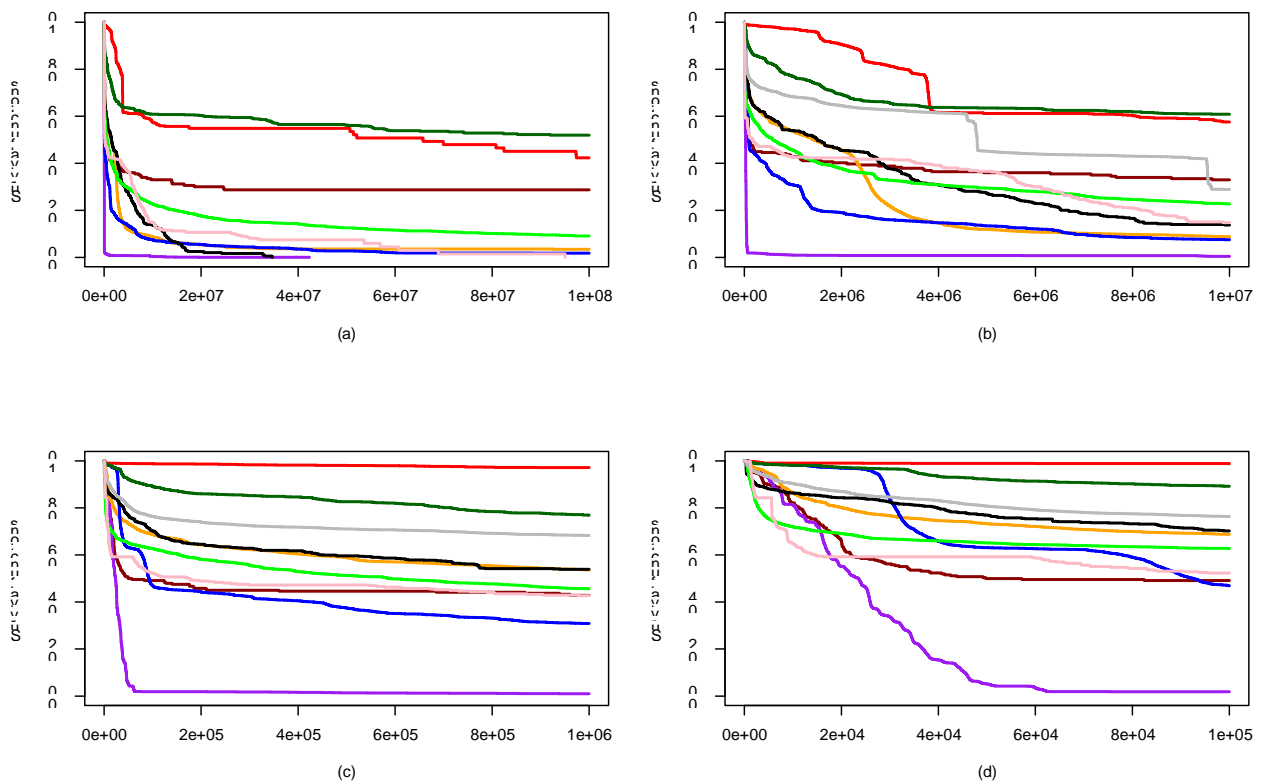


Рис.6 Оценки функций распределений времени успешного завершения прикладных задач в выбранных областях знаний

На Рис. 5 представлены: а) прикладные задачи, чье процессорное время исполнения, не превышающих  $1 \cdot 10^8$  сек; (b) - не превышающее  $1 \cdot 10^7$ ; (c) - не превышающие  $1 \cdot 10^6$ ; (d) - не превышающих  $1 \cdot 10^5$ .

**Заключение.** В работе рассмотрен подход к реализации моделей обучения системы планирования ресурсов гибридных суперкомпьютерных кластеров, работающих в режиме центров коллективного пользования, с использованием моделей внимания. Применяемый подход позволяет: а) использовать данные, полученных непосредственно из суперкомпьютерной вычислительной среды в форме лог-файлов работы планировщика заданий, б) сравнивать информацию, носителями которой выступают сгенерированные модели прогноза времени исполнения заданий и вычислительные процессы их непосредственной реализации, в) планировать целенаправленные действия по выделению и реконфигурации вычислительных ресурсов, г) получать объяснения полученных результатов. Предложенный подход позволяет преодолеть ключевые ограничений традиционных моделей машинного обучения, таких как зависимость от больших объёмов данных и сложность адаптации в условиях динамично меняющихся входных признаков.

## REFERENCES

1. Wang P., Li Y., Reddy C.K. Machine learning for survival analysis: A survey // *ACM Computing Surveys*.2019, Vol. 51, P. 1–36.
2. Cox D.R. Regression models and life-tables // *Journal of the Royal Statistical Society, Series B (Methodological)*. 1972, Vol. 34, P. 187–220.
3. Burkart N., Huber M.F. A survey on the explainability of supervised machine learning // *Journal of Artificial Intelligence Research*. 2021; 70:245-317.
4. Sahakyan M., Aung Z., Rahwan T. Explainable artificial intelligence for tabular data: A survey // *IEEE Access*. 2021, Vol. 9, P. 135392-135422.
5. Ribeiro M.T., Singh S., Guestrin C. Why should I trust you? Explaining the predictions of any classifier // *arXiv:1602.04938*, 2016.
6. Beran R. Nonparametric regression with randomly censored survival data // Technical report, University of California, Berkeley, 1981.
7. Kovalev M.S., Utkin L.V., Kasimov E.M. SurvLIME: A method for explaining machine learning survival models // *Knowledge-Based Systems*. 2020, Vol. 203, P. 106164.
8. Utkin L.V., Kovalev M.S., Kasimov E.M. An explanation method for black-box machine learning survival models using the Chebyshev distance // *Artificial Intelligence and Natural Language. AINL 2020*, Springer, Cham, Communications in Computer and Information Science.2020, Vol. 1292, P. 62-74.
9. Kovalev M.S., Utkin L.V. A robust algorithm for explaining unreliable machine learning survival models using the Kolmogorov-Smirnov bounds // *Neural Networks*. 2020, Vol. 132, P. 1-18.
10. Agarwal R., Frosst N., Zhang X., Caruana R., Hinton G.E. Neural Additive Models: Interpretable Machine Learning with Neural Nets. *arXiv:2004.13912*, 2020.
11. Utkin L.V., Satyukov E.D., Konstantinov A.V. SurvNAM: The machine learning survival model explanation // *Neural Networks*. 2022, Vol. 147, P. 81–102.
12. Utkin L.V., Eremenko D.Y., Konstantinov A.V. SurvBeX: An explanation method of the machine learning survival models based on the Beran estimator // *arXiv:2308.03730*, 2023.
13. Utkin, L.V., Kirpichenko, S.R., Konstantinov, A.V., Zaborovsky, V.S., Orlov, I.N., Verbova, N.M. (2026). A Probabilistic Concept-Based Learning Method Within the Framework of Survival Analysis. In: Kovalev, S., Kotenko, I., Sukhanov, A. (eds) *Proceedings of the Ninth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’25)*, Volume 2. IITI 2025. Lecture Notes in Networks and Systems, vol 1763. Springer, Cham. [https://doi.org/10.1007/978-3-032-13612-1\\_10](https://doi.org/10.1007/978-3-032-13612-1_10)

14. Zaborovsky, V.S., Mulyukha, V.A., Veselov, V.F. (2026). A Turning Point in Computer Science: Generative Large Language Models Aren't All You Need. In: Kovalev, S., Kotenko, I., Sukhanov, A. (eds) Proceedings of the Ninth International Scientific Conference "Intelligent Information Technologies for Industry" (ITI'25), Volume 2. ITI 2025. Lecture Notes in Networks and Systems, vol 1763. Springer, Cham. [https://doi.org/10.1007/978-3-032-13612-1\\_52](https://doi.org/10.1007/978-3-032-13612-1_52).

**Уткин Л.В.,**

Техникағылымдарының докторы, профессор,  
[vlad2tu@yandex.ru](mailto:vlad2tu@yandex.ru)<sup>1</sup>

**Заборовский В.С.,**

Техникағылымдарының докторы, профессор,  
[vlad2tu@yandex.ru](mailto:vlad2tu@yandex.ru)<sup>1</sup>

*Ұлы Петр атындағы Санкт-Петербург политехникалық университеті,  
195251, Санкт-Петербург, Политехническая көш., 291<sup>1</sup>*

## ҰЖЫМДЫҚ ПАЙДАЛАНУҒА АРНАЛҒАН ОРТАЛЫҚТАР РЕТІНДЕГІ ГИБРИДТІ СУПЕРКОМПЬЮТЕРЛЕРДІҢ МАШИНАЛЫҚ ОҚЫТУ МӘСЕЛЕЛЕРІ ТУРАЛЫ

***Аңдатпа.** Есептеу технологияларының қарқынды дамуы және амбебап генеративті трансформер-модельдердің пайда болуы зерттеушілердің алдына маңызды сұрақ қояды: қандай жағдайларда және қандай міндеттер үшін формалды бағдарламалау тілдері арқылы жазылған қолданбалы есептеу бағдарламаларын алдын ала оқытылған мультимодальді үлкен тілдік модельдерге арналған промпт-сұраныстармен алмастыруға болады? Мақалада «физикалық құбылыс ақпараттықты туындатады, ал ақпараттық құбылыс есептеуді қалыптастырады» деген абдуктивтік тұжырымға сүйене отырып, ұжымдық пайдалануға арналған орталықтар режимінде жұмыс істейтін гибриді суперкомпьютерлік кластерлердің өнімділігіне машиналық оқыту үдерістерінің әсер ету механизмі қарастырылады. Зерттеу нәтижелері Жозеф Луи Лагранж ұсынған «Дауласпайық, есептеп көрейік» деген тұжырымның өзектілігін растайды.*

***Түйінді сөздер:** суперкомпьютерлер, машиналық оқыту, өмір сүру моделі, экзо-интеллект, түсіндірме интеллект, дәлелдеу, ықтималдық.*

**Utkin L.V.,**

Doctor of Technical Sciences, professor,  
vlad2tu@yandex.ru<sup>1</sup>

**Zaborobsky V.S.,**

Doctor of Technical Sciences, professor,  
vlad2tu@yandex.ru<sup>1</sup>

*Saint Petersburg Polytechnic University*

*Peter the Great*

*195251, Saint Petersburg, Politekhnicheskaya St., 291<sup>1</sup>*

## **ON THE PROBLEM OF MACHINE LEARNING FOR HYBRID SUPERCOMPUTERS AS SHARED USE CENTERS**

***Abstract/** The rapid development of computing technologies and the creation of universal generative models-transformers poses a fundamental question for researchers: under what conditions and for what tasks can computational applications written using formal languages be replaced with prompt queries to pre-trained multimodal large-scale language models.*

*Drawing on the abductive inference that the "physical" gives rise to the "informational," and that "the informational gives rise to the computational," this article examines the mechanism by which machine learning processes influence the performance of hybrid supercomputer clusters operating in shared-use center mode, confirming the relevance of J.-L. Lagrange's thesis: "Let's not argue, let's calculate."*

***Keywords:** supercomputers, machine learning, survival model, exo-intelligence, explanatory intelligence, proof, probability.*