

ЗЫРЯНОВ Дмитрий Александрович,
студент,
Санкт-Петербургский государственный
университет аэрокосмического приборостроения,
e-mail: zyrkdima@gmail.com

АНАЛИЗ ПРИГОДНОСТИ РАЗЛИЧНЫХ АРХИТЕКТУР НЕЙРОННЫХ СЕТЕЙ НА СИНТЕТИЧЕСКОМ НАБОРЕ ДАННЫХ ДЛЯ ЗАДАЧИ ВЫБОРА МАРШРУТА

Аннотация. В статье рассматривается решение задачи выбора наиболее подходящего маршрута движения беспилотных транспортных средств среди нескольких альтернатив. При выборе оцениваются параметры маршрутов (протяженность, средняя скорость трафика, проходимость, погода, количество заправочных станций), а также параметры точки старта (количество заправочных станций, стоимость парковки, погода) для возможности принятия решения о приостановке движения. Для анализа формируется синтетический набор данных с явной корреляцией параметров. Применимость датасета проверяется на линейных моделях с различным количеством входных параметров. Рассматриваются различные архитектуры нейронных сетей: линейные, рекуррентные (RNN), нейронные сети с долгой краткосрочной памятью (LSTM), трансформеры. Приведен анализ работы моделей. Определяются следующие преимущества и недостатки рассматриваемых архитектур для задачи: LSTM и RNN наименее подвержены переобучению на наборе данных маршрутов; затраты на время обучения у RNN меньше, чем у других рассмотренных моделей; скорость вычисления у LSTM и RNN выше, чем у transformer (при одинаковом количестве параметров); при малом количестве параметров точность модели transformer выше. Сделан вывод о том, что transformer лучше справляется с задачей по выявлению связей между параметрами маршрута для выбора лучшего варианта, однако необходимо следить за переобучением, также благодаря тому, что Transformer требует меньшее количество параметров для стабильной работы, это может компенсировать проигрыш во времени расчета в сравнении с RNN и LSTM.

Ключевые слова: нейронная сеть, трансформер, рекуррентная сеть, LSTM, RNN, беспилотный аппарат, поиск маршрута, управление, данные.

ZYRYANOV Dmitry Alexandrovich,
student,
Saint-Petersburg State University of Aerospace Instrumentation

ANALYSIS OF THE SUITABILITY OF VARIOUS NEURAL NETWORK ARCHITECTURES ON A SYNTHETIC DATASET FOR THE ROUTE SELECTION TASK

Annotation. The article considers the solution to the problem of choosing the most suitable route for unmanned vehicles among several alternatives. When choosing, the route parameters (length, average traffic speed, passability, weather, number of gas stations) are evaluated, as well as the parameters of the starting point (number of gas stations, parking cost, weather) for the possibility of making a decision to suspend traffic. A synthetic data set with an explicit correlation of parameters is formed for the analysis. The applicability of the dataset is tested on linear models with a different number of input parameters. Various architectures of neural networks are considered: linear, recurrent (RNN), neural networks with long short-term memory (LSTM), transformers. An analysis of the operation of the models is given. The following advantages and disadvantages of the considered architectures for the task are determined: LSTM and RNN are least susceptible to retraining on a set of route data; RNN training time costs are lower than for other models considered; The calculation speed of LSTM and RNN is higher than that of transformer (with the same number of parameters); with a small number of parameters, the accuracy of the transformer model is higher. It is concluded that transformer copes better with the task of identifying relationships between route parameters in order to choose the best option, however, it is necessary to monitor retraining, also due to the fact that Transformer requires fewer parameters for stable operation, this can compensate for the loss in calculation time compared to RNN and LSTM.

Key words: *neural network, transformer, recompetitive network, LSTM, RNN, unmanned vehicle, route search, control, data.*

Одной из задач беспилотного транспорта является выбор наиболее подходящего маршрута движения среди нескольких альтернатив. Данная задача является общей для различных навигационных систем, систем принятия решения и оптимизации. Различные варианты решения существуют для морских [1], воздушных [2], наземных [3] транспортных систем. Важно отметить, что критерии выбора могут неочевидным образом коррелировать с входными параметрами маршрутов, при этом определение оптимального без уменьшения количества параметров становится неразрешимой задачей [1]. Хорошие результаты в решении задач маршрутизации показывают различные алгоритмы искусственного интеллекта: генетические алгоритмы [4], модели машинного обучения с подкреплением [5], муравьиный алгоритм [2], а также нейронные сети [6]. Поэтому для задачи выбора маршрута предлагается применение нейронных сетей, которые наиболее приспособлены для поиска скрытых связей между данными. Однако важно опреде-

лить архитектуру, наиболее подходящую для данной задачи.

Для анализа моделей сформируем синтетический датасет. Набор данных создадим случайным образом с указанием явной зависимости между параметрами (для проверки работы моделей). Входные данные были выбраны следующим образом:

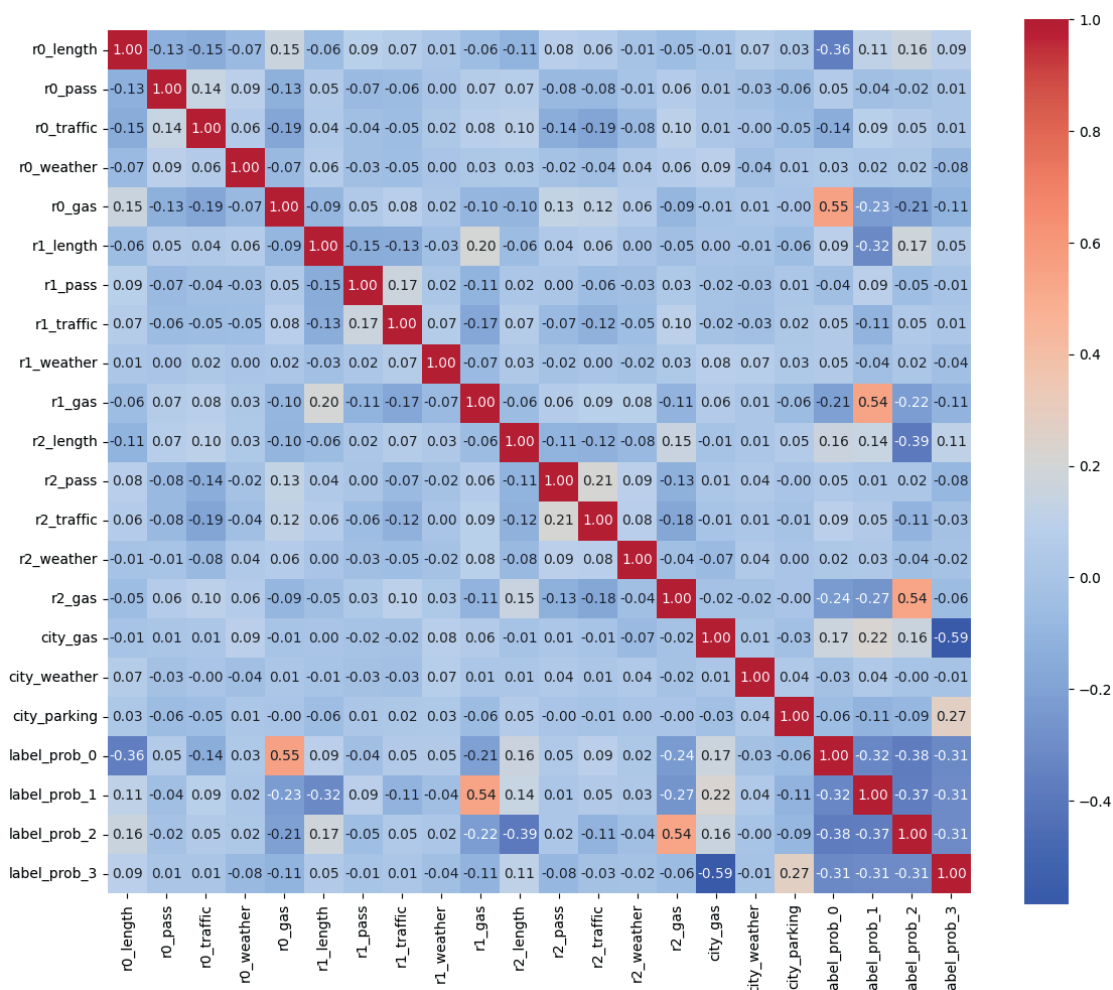
1) Параметры маршрутов (для анализа сформировано по 3 альтернативных маршрута): длина маршрута, проходимость (в виде коэффициента от 0 до 1), средняя скорость трафика, погодные условия, количество заправочных станций на маршруте;

2) Параметры города начала пути: количество заправочных станций, погода, стоимость стоянки.

Выходные данные - распределение вероятностей выбора одного из маршрутов или города (стоянка без продолжения следования).

Для анализа пригодности была сформирована матрица корреляции (рис. 1).

Рисунок 1. Матрица корреляции между данными
Данная матрица показывает, что между данными существует определенная связь, а значит они могут быть использованы для дальнейшей оценки моделей [7].



Для начала необходимо определить, сможет ли система работать с постоянным количеством входных параметров. Для этого сравним 3 полностью связанные модели: с 1 маршрутом, с 2 маршрутами, с 3 маршрутами. При этом выбираются маршруты с наименьшей длиной пути – оставшиеся отсеиваются. Это определяется из требований ко всем навигационным алгоритмам по поиску кратчайших маршрутов.

Модели имеют одинаковую архитектуру: 3 последовательных линейных слоя с батч-нормализацией, dropout и функцией активации LeakyReLU, выходной линейный слой с функцией активации softmax.

Обучение длится 1300 эпох, применяется

оптимайзер Adam и функция ошибки – Cross-Entropy.

Для моделей введены 2 оценки точности:

1) правильность выбора конкретного маршрута или остановки в городе;

2) правильность выбора любого из маршрутов или остановки в городе.

Последняя оценка покажет, будут ли дополнительные маршруты влиять на результаты работы моделей.

Все данные делятся на 2 выборки:

1) тренировочная – 80% данных;

2) тестовая – 20% данных.

Оценка производится на тестовой выборке – на данных, которые модели никогда не видели.

Рисунок 2. График обучения модели с 1 входным маршрутом.

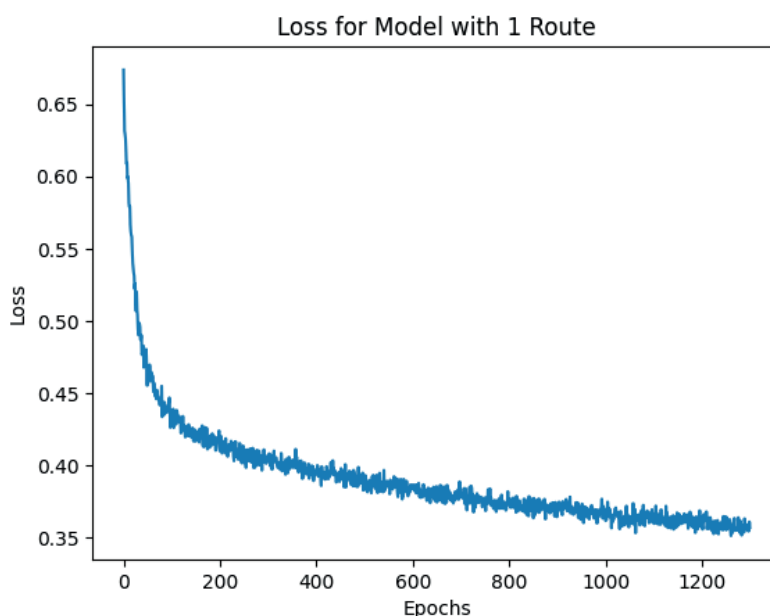


Рисунок 3. График обучения модели с 2 входными маршрутами.

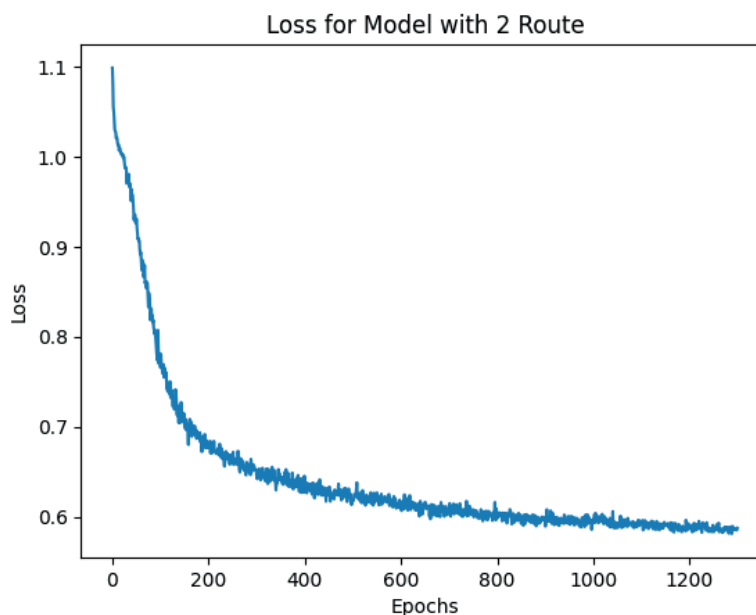
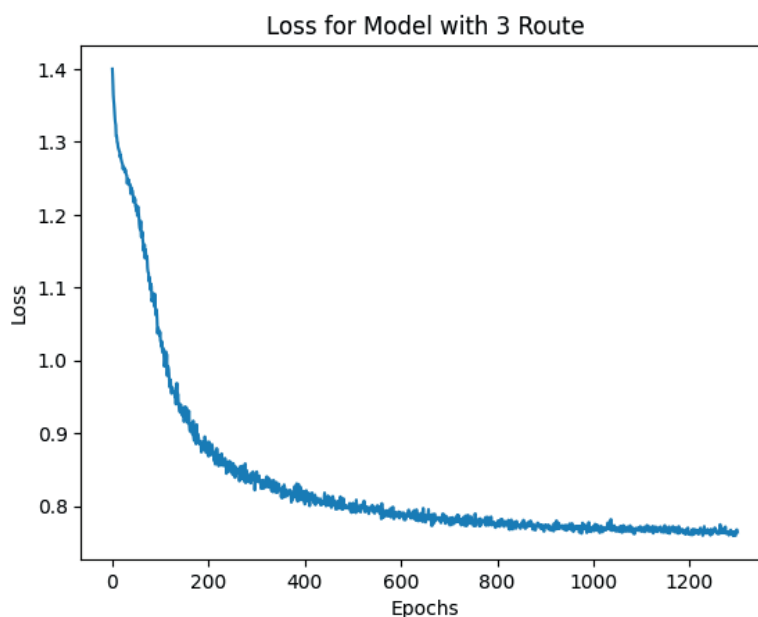


Рисунок 4. График обучения модели с 3 входными маршрутами.



Из графиков видно, что модели хорошо обучены и достигли стагнации.

Результаты оценки точности:

Модель с 1 входным маршрутом:

- точность выбора маршрута или остановки: 82.5%;

- точность выбора любого маршрута или остановки: 82.5%.

Модель с 2 входными маршрутами:

- точность выбора маршрута или остановки: 84.5%;

- точность выбора любого маршрута или остановки: 86.5%.

Модель с 3 входными маршрутами:

- точность выбора маршрута или остановки: 81%;

- точность выбора любого маршрута или остановки: 87.5%.

Данный анализ показывает, что при введении в модель информации об альтернативных маршрутах меняется точность определения действия «остаться в городе или продолжить движение». При этом каждый следующий маршрут будет вносить меньше влияния в данную оценку. Следовательно, необходимо рассматривать модели, принимающие переменное количество маршрутов для различных обстоятельств [8]. Также можно применять постоянное количество альтернативных маршрутов, но в этом случае если

маршрутов будет меньше, придется заполнять пропуски 0, из-за чего может быть неверная корреляция. Другой альтернативой является кодирование всех маршрутов в одном. Однако в этом случае необходимо рассмотреть преобразования данных: например, как объединить информацию о погоде на маршрутах, или о длине пути. Т.к., во-первых, логических способов объединить данные нет, во-вторых, в будущем данные будут добавляться для улучшения работы системы. И самое главное, мы не сможем выделить один маршрут из объединения, поэтому данный вариант неприменим.

Рассмотрим основные виды современных моделей [9]:

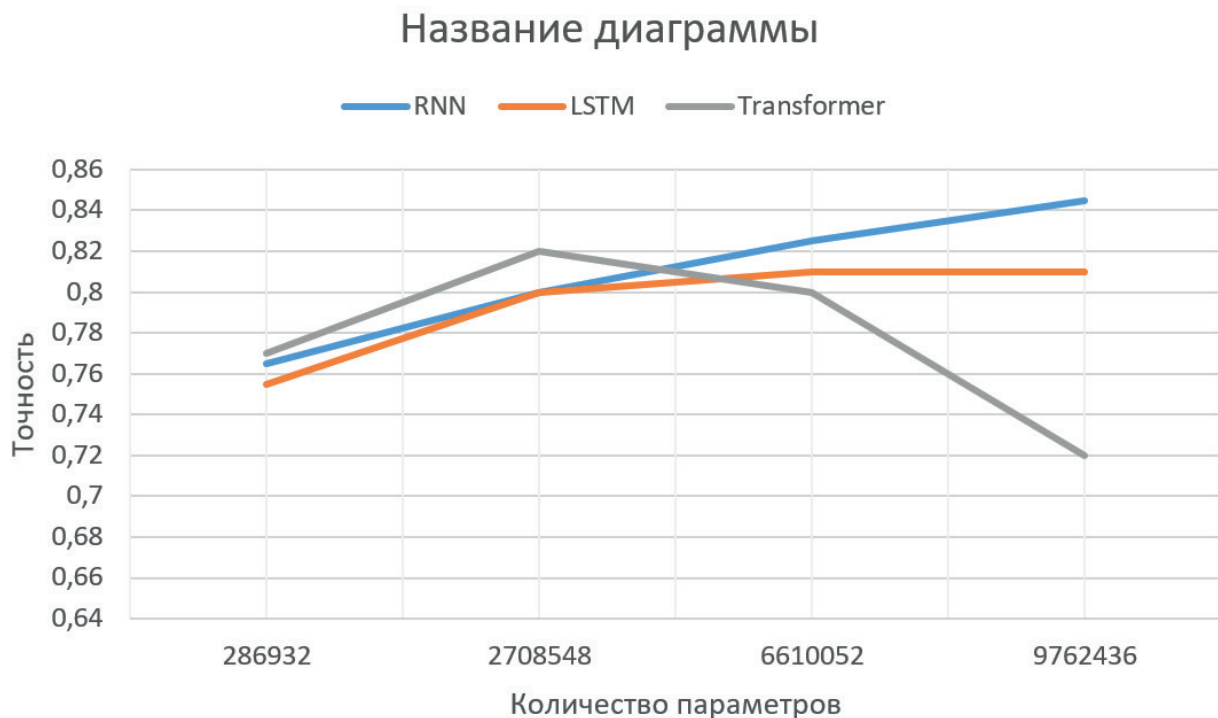
- 1) RNN – рекуррентная нейронная сеть. В исследовании будет применяться ее вариант GRU, в котором решена проблема взрывных и затухающих градиентов.

- 2) LSTM.

- 3) Transformer [10].

Для сравнения моделей между собой необходимо подобрать оптимальное количество параметров для каждой из моделей. Для этого будем сравнивать по 4 модели с различным количеством параметров. Оцениваться будут точность и количество параметров. Последний критерий будет прямым образом влиять на скорость расчетов модели и на потребление памяти.

Рисунок 5. Зависимости точности моделей от количества параметров.



На рисунке 5 можно увидеть, что с данным набором данных лучше всего справляется RNN (GRU). При этом стоит отметить, что Transformer при малом количестве параметров имеет хорошую точность, но при большем – наоборот точность начинает снижаться. Это обусловлено быстрым переобучением модели трансформера на приведенном наборе данных.

Важно отметить, что благодаря простоте внутреннего алгоритма GRU он рассчитывается быстрее, чем модель Transformer с тем же количеством параметров.

Сформируем следующие выводы:

- LSTM и RNN наименее подвержены переобучению на наборе данных маршрутов;

- затраты на время обучения у RNN меньше, чем у других рассмотренных моделей;

- скорость вычисления у LSTM и RNN выше, чем у transformer (при одинаковом количестве параметров);

- при малом количестве параметров точность модели Transformer выше.

Таким образом, модель Transformer лучше справляется с задачей по выявлению связей между параметрами маршрута для выбора лучшего варианта, однако необходимо следить за переобучением. Благодаря тому, что Transformer требует меньшее количество параметров для стабильной работы, это может компенсировать проигрыш во времени расчета в сравнении с RNN и LSTM.

Список литературы:

[1] Андреева Е. В. Многокритериальный подход в задаче выбора оптимальных маршрутов в акватории Северного морского пути / Е. В. Андреева // Вестник государственного университета морского и речного флота им. адмирала С.О. Макарова. – 2021. – Т. 13, № 3. – С. 399-408. – DOI 10.21821/2309-5180-2021-13-3-399-408.

[2] Задача выбора оптимального маршрута беспилотного летательного аппарата при противодействующих, географических и погодных ограничениях в области полёта / К. Е. Лавренов, А. П. Шевелева, О. Т. Романов, М. Н. Машкин // Информационные технологии в проектировании и производстве. – 2021. – № 4(184). – С. 22-28.

[3] Новиков Л. В. Формирование маршрута с учетом коэффициента проходимости дорог для наземной группы поиска отделяемых частей ракет космического назначения и их фрагментов / Л. В. Новиков, О. Л. Шестопалова // Вопросы технических и физико-математических наук в свете современных исследований : Сборник статей по материалам XXXIX международной научно-практической конференции, Новосибирск, 26 мая 2021 года / ООО «СибАК». Том 5 (31). – Новосибирск: Общество с ограниченной ответственностью «Сибирская академическая книга», 2021. – С. 32-37.

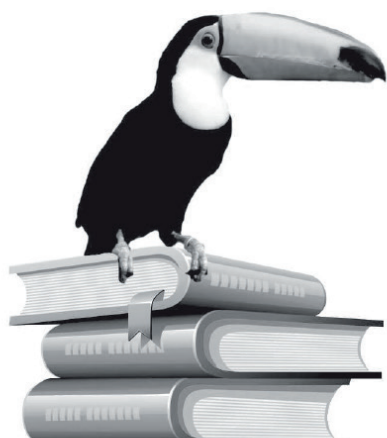
- [4] Зырянов Д. А. Исследование применения генетических алгоритмов в задачах планирования маршрута в сравнении с эвристическими / Д. А. Зырянов, Е. М. Бади́ка, А. П. Бобрышов // Завалишинские чтения 25 : сборник докладов XX Международной конференции по электромеханике и робототехнике, Санкт-Петербург, 15–16 апреля 2025 года. – Санкт-Петербург: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2025. – С. 153-157.
- [5] Шмыглев Д. Н. Модель обучения с подкреплением для оптимизации автопарка предприятия / Д. Н. Шмыглев, В. А. Судаков // Препринты ИПМ им. М.В. Келдыша. – 2024. – № 39. – С. 1-13. – DOI 10.20948/prepr-2024-39.
- [6] Якимов, М. А. Анализ алгоритмов поиска маршрута в транспортной сети / М. А. Якимов, К. В. Операйло // Модели и методы повышения эффективности инновационных исследований : Сборник статей по итогам Международной научно-практической конференции, Воронеж, 04 июня 2021 года. – Стерлитамак: Общество с ограниченной ответственностью «Агентство международных исследований», 2021. – С. 176-184.
- [7] Траск Э. Грокаяем глубокое обучение. СПб.: Питер, 2019. 352 с.
- [8] Качалов, О. И. Сравнение точности прогнозирования стоимости акций компании с использованием свёрточной нейронной сети (CNN), простой рекуррентной нейронной сети (SIMPLE RNN) и рекуррентной нейронной сети на базе долгой краткосрочной памяти (LSTM) / О. И. Качалов, А. Н. Миронов, А. М. Володина // ИТ-Стандарт. – 2018. – № 4(17). – С. 49-61.
- [9] Булатов, А. С. Обзор архитектуры рекуррентного Трансформера в контексте нейронных сетей с памятью / А. С. Булатов, Ю. М. Куратов, М. С. Бурцев // ТРУДЫ МФТИ. Труды Московского физико-технического института (национального исследовательского университета). – 2024. – Т. 16, № 4(64). – С. 5-20.
- [10] Патент № 2841111 С1 Российская Федерация, МПК В60W 60/00, G06N 3/02, G06V 10/82. Способ управления бортовыми системами беспилотных транспортных средств при помощи нейронных сетей на основе архитектуры трансформеров : заявл. 28.06.2024 : опубл. 02.06.2025 / А. А. А. Э. Карим, М. А. А. М. Хегази, Б. Рашид ; заявитель Автономная некоммерческая организация высшего образования «Университет Иннополис».

References:

- [1] Andreeva E. V. Multicriteria Approach to Selecting Optimal Routes in the Waters of the Northern Sea Route / E. V. Andreeva // Bulletin of the Admiral S. O. Makarov State University of Maritime and Inland Shipping. - 2021. - Vol. 13, No. 3. - Pp. 399-408. - DOI 10.21821/2309-5180-2021-13-3-399-408.
- [2] The Problem of Selecting the Optimal Route of an Unmanned Aerial Vehicle Subject to Counteracting, Geographical, and Weather Restrictions in the Flight Area / K. E. Lavrenov, A. P. Sheveleva, O. T. Romanov, M. N. Mashkin // Information Technologies in Design and Production. - 2021. - No. 4 (184). – P. 22-28.
- [3] Novikov L. V. Formation of a route taking into account the road passability coefficient for a ground group searching for detachable parts of space rockets and their fragments / L. V. Novikov, O. L. Shestopalova // Issues of technical and physical-mathematical sciences in light of modern research: Collection of articles based on the materials of the XXXIX international scientific and practical conference, Novosibirsk, May 26, 2021 / SibAK LLC. Volume 5 (31). - Novosibirsk: Limited Liability Company «Siberian Academic Book», 2021. - P. 32-37.
- [4] Zyryanov D. A. Study of the application of genetic algorithms in route planning problems in comparison with heuristic ones / D. A. Zyryanov, E. M. Badika, A. P. Bobryshov // Zavalishin Readings 25: collection of reports of the XX International Conference on Electromechanics and Robotics, St. Petersburg, April 15–16, 2025. – St. Petersburg: St. Petersburg State University of Aerospace Instrumentation, 2025. – Pp. 153–157.
- [5] Shmyglev D. N. Reinforcement learning model for optimizing an enterprise's vehicle fleet / D. N. Shmyglev, V. A. Sudakov // Preprints of the Keldysh Institute of Applied Mathematics. – 2024. – No. 39. – Pp. – DOI 10.20948/prepr-2024-39.
- [6] Yakimov, M. A. Analysis of Route Search Algorithms in a Transport Network / M. A. Yakimov, K. V. Operaylo // Models and Methods for Improving the Efficiency of Innovative Research: Collection of Articles Following the Results of the International Scientific and Practical Conference, Voronezh, June 4, 2021. – Sterlitamak: International Research Agency, Limited Liability Company, 2021. – Pp. 176-184.
- [7] Trask, E. Grokking Deep Learning. St. Petersburg: Piter, 2019. 352 p.
- [8] Kachalov, O. I. Comparison of the Accuracy of Forecasting the Price of a Company's Stocks Using a Convolutional Neural Network (CNN), a Simple Recurrent Neural Network (SIMPLE RNN), and a Recurrent Neural Network Based on Long Short-Term Memory (LSTM) / O. I. Kachalov, A. N. Mironov, A. M. Volodina // IT-Standard. - 2018. - No. 4 (17). - Pp. 49-61.
- [9] Bulatov, A. S. Review of the Architecture of a Recurrent Transformer in the Context of Neural Networks with Memory / A. S. Bulatov, Yu. M. Kuratov, M. S. Burtsev // MIPT PROCEEDINGS. Proceedings of the Moscow Institute of Physics and Technology (National Research University). - 2024. - Vol. 16, No. 4

(64). - Pp. 5-20.

[10] Patent No. 2841111 C1 Russian Federation, IPC B60W 60/00, G06N 3/02, G06V 10/82. Method for controlling on-board systems of unmanned vehicles using neural networks based on transformer architecture : declared 28.06.2024 : published 02.06.2025 / A. A. A. E. Karim, M. A. A. M. Hegazy, B. Rashid ; applicant Autonomous Non-Commercial Organization of Higher Education «Innopolis University».



ЮРКОМПАНИ

www.law-books.ru

Юридическое издательство
«ЮРКОМПАНИ»

Издание учебников,
учебных и методических
пособий, монографий,
научных статей.

Профессионально.

В максимально
короткие сроки.

Размещаем
в РИНЦ, E-Library.