

СРАВНЕНИЕ ЧИСЛА MLP-БЛОКОВ В MLP-MIXER

Монтик Н. С., ст. преп.

*Брестский государственный технический университет,
г. Брест, Республика Беларусь*

Реферат. В статье рассмотрено сравнение различного числа MLP-блоков в архитектуре MLP-Mixer. Выбор числа блоков в данной архитектуре имеет важное значение для точности классификации и скорости обучения нейронной сети. В качестве датасета использовался CIFAR-10.

Ключевые слова: глубокое обучение, нейронная сеть, MLP-Mixer, CIFAR-10.

CIFAR-10 – это датасет, используемый в задачах машинного обучения, особенно в области компьютерного зрения и классификации изображений. CIFAR расшифровывается как Canadian Institute For Advanced Research, а число 10 указывает на количество классов объектов, представленных в датасете.

В датасете содержится 60,000 цветных изображений размером 32x32 пикселя: 50,000 изображений предназначены для обучения модели, а оставшиеся 10,000 используются для тестирования. Все изображения распределены по 10 классам (самолёт, автомобиль, птица, кошка, олень, собака, лягушка, лошадь, корабль и грузовик), каждый из которых содержит по 6,000 изображений. Изображения представлены в формате RGB, то есть имеют три цветовых канала: красный, зелёный и синий.

MLP-Mixer (MLP – multilayer perceptron) – это архитектура нейронной сети, предложенная исследователями из Google Research в 2021 году [1]. Она предназначена для решения задач компьютерного зрения и представляет собой альтернативу популярным свёрточным нейронным сетям и трансформерам. Главная особенность MLP-Mixer заключается в том, что она полностью основана на многослойных перцептронах без использования свёрток или механизмов внимания.

MLP-Mixer достигает высокой точности благодаря комбинации линейных преобразований, нелинейных активаций и нормализации данных. Входные данные в MLP-Mixer сначала разбиваются на патчи, затем каждый патч преобразуется в вектор фиксированной длины с помощью линейного преобразования. После этого данные подаются в последовательность MLP-блоков. В каждом блоке сначала выполняется операция смешивание каналов, которая работает с признаками внутри каждого патча независимо. Эта операция выявляет зависимости между различными каналами данных, например, между цветовыми компонентами или признаками, извлечёнными из патчей. Затем выполняется операция смешивания пространственных отношений, которая учитывает взаимосвязи между патчами в пространстве. Это позволяет модели понять, как различные части изображения связаны друг с другом.

Каждый MLP-блок включает в себя нормализацию данных и нелинейную функцию активации (например, GELU). Нормализация применяется до и после каждой операции смешивания, чтобы стабилизировать обучение и улучшить сходимость модели. После прохождения через несколько таких блоков выходные данные подаются в полносвязный слой для получения окончательного результата, например, метки класса в задаче классификации изображений. Такая структура делает MLP-Mixer простым в реализации, но при этом достаточно мощным для решения сложных задач.

Для реализации поставленной задачи использовался язык программирования Python, библиотека pytorch. Инициализация весов стандартная, все алгоритмы оптимизации используют стандартные параметры, значение seed установлено равным 42. В качестве оптимизатора выбран Adam.

Размер изображения выбран в соответствии с датасетом: 32x32, 3 канала. Разбиение происходит на 64 патча размером 4x4 пикселя. Каждый патч преобразуется в вектор размерности 128. Как и в оригинальном MLP-Mixer [1], используется skip connection на обоих блоках нейронной сети: значение перед выполнением нормализации слоя передаётся на следующий блок нейронной сети.

Размер батча установлен равным 32. Обучение выполнялось в течение 10 эпох на следующей ЭВМ: процессор Intel Core i7-12700KF, оперативная память 2x16GB 4000MHz CL16, видеокарта

Nvidia GeForce RTX 3080 12GB.

В качестве метрик оценки производительности использовались длительность обучения и точность. Результаты сравнения представлены в таблице 1.

Таблица 1 – Сравнение точности и скорости обучения

Метрика	Число блоков					
	2	4	6	8	10	12
Точность, %	64,94	67,53	67,91	69,21	68,1	67,51
Длительность обучения, сек.	149,36	194,38	249,77	296,03	333,17	364,29

Как показывают результаты, максимальная точность достигается при использовании 8 блоков MLP-Mixer. Скорость обучения при этом остаётся на приемлемом уровне.

Таким образом, оптимальное число блоков для данной задачи составляет 8, что обеспечивает баланс между точностью и временем обучения.

Список использованных источников

1. MLP-Mixer: An all-MLP Architecture for Vision [Electronic resource] / Ilya T., Neil H., Alexander K., Lucas B., Xiaohua Z., Thomas U., Jessica Y., Andreas S., Daniel K., Jakob U., Mario L., Alexey D. // Cornell University. – 2021. – Mode of access: <https://arxiv.org/pdf/2105.01601/>. – Date of access: 07.03.2024.

УДК 004.05

МЕТОДЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ ПРОЦЕССА РАЗРАБОТКИ

Знатнов Г. Д., студ., Каргина Л. Р., д.э.н., проф.

Российский университет транспорта, г. Москва, Российская Федерация

Реферат. В статье рассмотрены теоретические основы оценки эффективности разработки, разобраны две методологии DORA и SPACE. Оценка эффективности разработки имеет важное значение, так как от этого напрямую зависят успехи любой IT-компаний, их финансовые показатели и бизнес-результаты.

Ключевые слова: эффективность разработки, производительность, разработка, эффективность, фреймворки, SPACE, DORA, метрики эффективности, IT-технологии.

Сфера IT-технологий является одной из самой крупных в современном бизнесе – по прогнозам компании Forrester общемировые расходы на IT-технологии в 2025 году поднимутся до значения 5,6 % и достигнут \$ 4,9 трлн. Важнейшими факторами роста являются активное внедрение программного обеспечения и IT-услуг во все сферы жизни, высокий спрос и развитие искусственного интеллекта, а также облачных технологий. IT-компании вкладывают и зарабатывают огромные деньги, именно поэтому поднимается спрос на любые методы оптимизации производительности бизнес-процессов – для получения наибольшей выручки. Так как разработчики являются важнейшим механизмом создания IT-продукта, инициативы по улучшению эффективности процесса разработки выходят на первый план.

Эффективность разработки заключается не в том, чтобы производить как можно больше программного кода, а в том насколько эффективно этот код был написан и насколько качественно и быстро он приводит к достижению бизнес-результатов компании. Однако в этом и заключается проблема: процесс разработки является очень комплексным и не может быть оценен количественно, в отличие от маркетинга или сферы продаж, где все можно измерить цифрами. Таким образом, в измерении эффективности разработки важно оценивать как количественные, так и качественные показатели.