

White Paper V3.2

Октябрь 2017

1. Введение
2. Платформа Neuromation
3. Фундаментальные основы
 - a. Нейронные сети и глубокое обучение
 - b. Искусственные данные
 - c. Распределенная обработка данных и экономика Neuromation
 - d. Модели машинного обучения
4. Практическое применение
5. Команда
6. Бизнес-модель Neuromation
7. Дорожная карта Neuromation
8. Продажа токенов
9. Экономика токена

Искусственный интеллект (ИИ) переживает новый расцвет и получает глобальное распространение в обществе. Исследователи и инженеры добились значительных успехов в решении практических проблем машинного обучения, особенно в таких прикладных сферах, как компьютерное зрение, распознавание речи, машинный перевод и принятие решений, индустриальная автоматизация и автономные машины. .

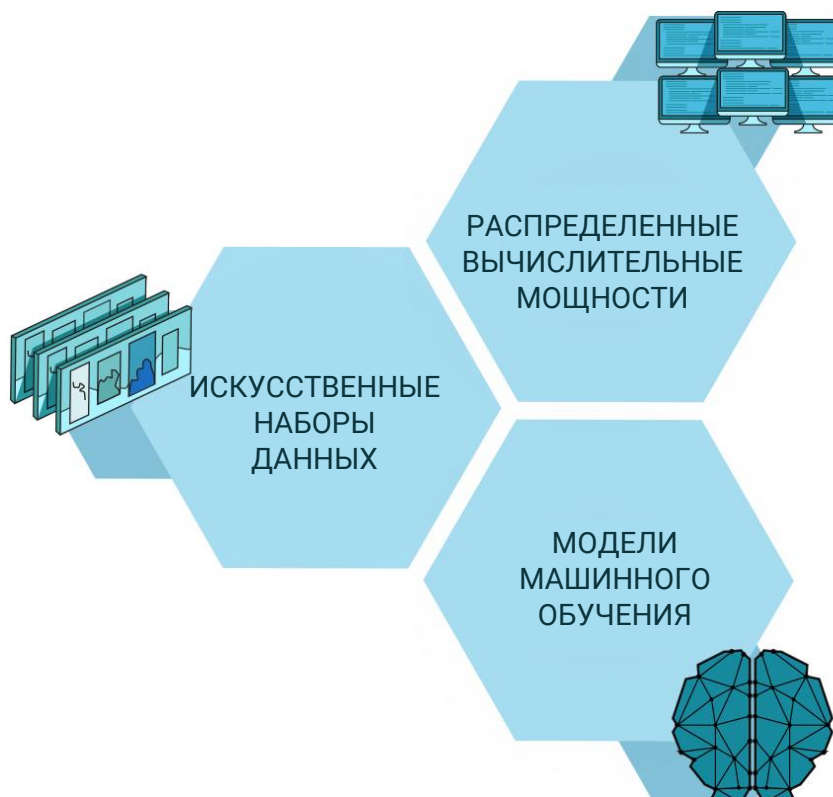
Благодаря развитию технологий в сочетании с прогрессом в исследованиях и инновациях нейросетевые алгоритмы быстро становится привычной частью деятельности любой компании. Такое развитие требует привлечения компетенций, вычислительных ресурсов, а также данных в достаточных объемах. И для индустрии машинного обучения проблемой является наличие наборов хорошо размеченных данных. На фоне возрастающих потоков неструктурированных, “сырых” данных, их преобразование в наборы данных с разметкой является ручным, чрезвычайно трудоемким и дорогостоящим процессом. Именно это ограничение является вызовом для широкого практического распространения глубокого обучения машинных алгоритмов на уровне доступной платформы с комфортным пользовательским интерфейсом. Пока это не сделано, у большинства компаний неизбежно возникают затруднения с автоматизацией методами машинного обучения. Из-за несоответствия в ресурсах и приоритете среди компаний нарастает разрыв между теми, кто начал переход на технологии глубокого обучения нейросетей и распределённых вычислений, и традиционными операторами, сохраняющими прежний технологический уклад. Стратегически, это поставит традиционные институты в крайне невыгодное положение в будущем.

Neuromation предлагает привлекательное решение, объединяющее рыночные ресурсы, научное сообщество, компании и частных лиц на интегрированной торговой площадке компонентов ИИ— **платформе Neuromation - добывающей знания**. Для решения наиболее актуальной проблемы индустрии Neuromation сосредотачивается на порождении наборов синтетических данных и эффективно тренирует с помощью таких данных модели нейросетей. Исследования компании позволили достигать результата работы машинных алгоритмов, равного по эффективности моделям нейросетей, натренированным на наборах реальных данных, размеченными вручную. Использование наборов синтетических данных в машинном обучении обеспечит дальнейшее снижение затрат на автоматизацию, ускорит и сделает широкое внедрение ИИ доступным для широкого круга операторов рынка в десятках секторов экономики.

Экосистема ИИ базируется на трех краеугольных камнях:

- **Наборы данных** – большие наборы структурированных (неструктурированных) данных, используемые в алгоритмах глубокого обучения, для обучения конкретной целевой функции.
- **Модели машинного обучения** – совокупность алгоритмов, предназначенных для обработки наборов данных.
- **Вычислительные мощности** – автономный сервер или ферма взаимосвязанных графических процессоров, предназначенных для обеспечения вычислений, необходимых для тренировки моделей ИИ с помощью специально созданных наборов данных.

В каждой из этих областей возникает свой уникальный набор проблем в отношении как возможностей, так и потребностей экосистемы ИИ. Neuromation предлагает новый элегантный подход к объединению спроса и предложения в каждой из этих областей на торговой площадке, сочетающий умения исследователей, инженеров, дизайнеров и связывающий их с компаниями коммерческого сектора.



Обзор

Мы создаем ключевой компонент зарождающейся экосистемы ИИ и планируем извлечь выгоду из преимуществ первопроходца: Платформа Neuromation обеспечивает биржевую площадку и среду, в которой участники могут как вносить свой вклад в создание модели ИИ, так и приобретать ее компоненты.

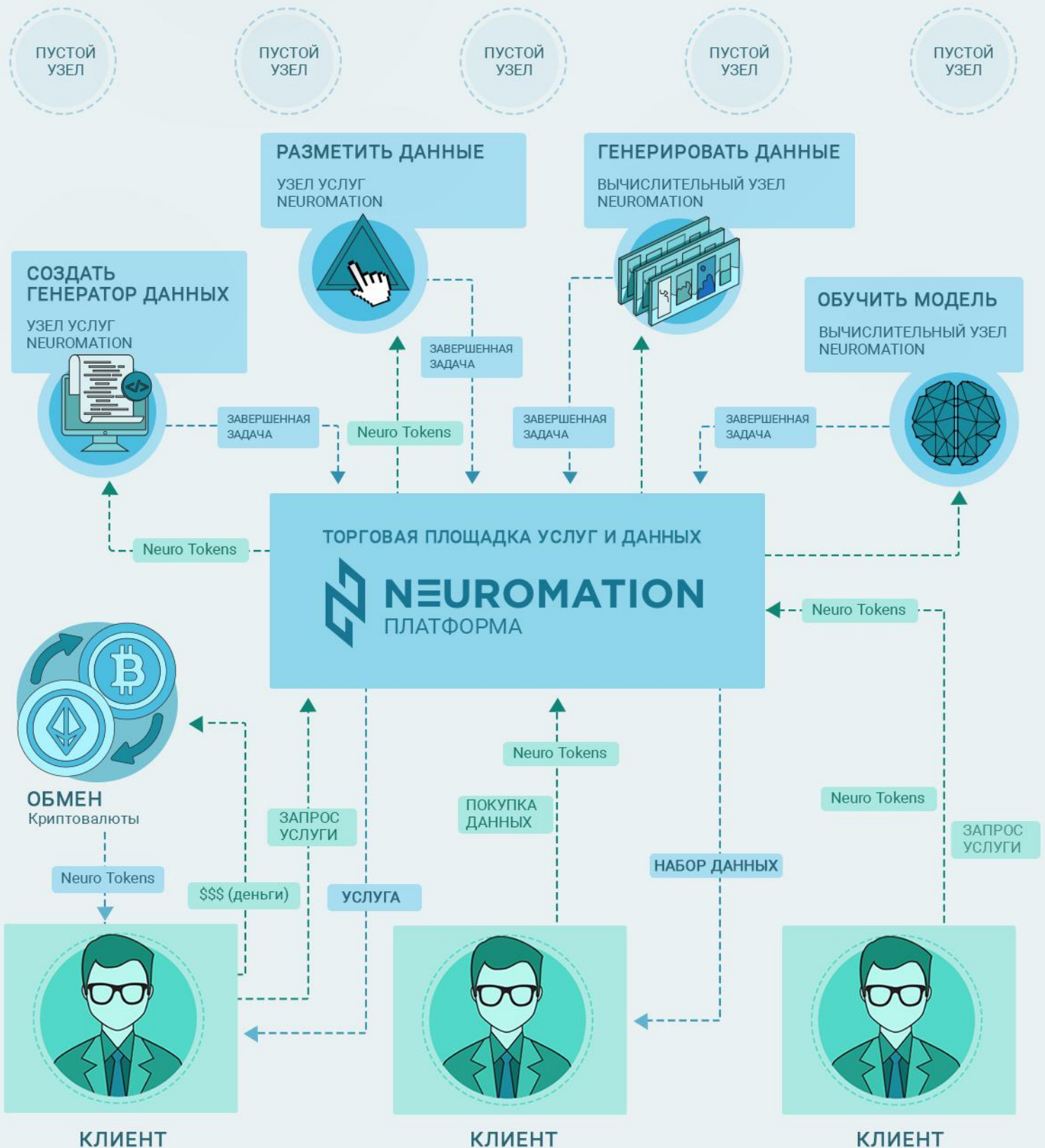
Платформа будет использовать распределенную обработку данных наряду с блокчейн-токенами proof-of-work для революционного преобразования процесса разработки моделей ИИ. Она объединит в одном месте все компоненты, необходимые для создания решений в области глубокого обучения. Поставщики услуг для платформы, коммерческие или частные, будут, в соответствии с "тремя краеугольными камнями" ИИ, указанными на предыдущем слайде, предоставлять определенные ресурсы для создания наборов искусственных данных, осуществления распределенных вычислительных услуг и разработки моделей машинного обучения. За каждую оказанную услугу (генерация наборов данных, обсчет моделей) или проданные данные (набор данных или генератор данных) полагается вознаграждение, выраженное в нашей валюте - Neurotoken(NTK). На платформе Neuromation будет использоваться аукционная модель, предусматривающая ведение клиентами переговоров с поставщиками услуг о ценах напрямую.

Представьте себе место, куда вы можете прийти и легко получить все необходимое для формирования потенциала ИИ. Поставщик создаст для вас генератор данных, затем группа узлов Neuromation Nodes использует генератор для быстрого создания обширного набора виртуальных данных. Далее, на основе этих данных, вы можете выбрать набор архитектур глубокого обучения для тренировки. Затем другая группа узлов Neuromation Nodes проведет тренировку в рекордные сроки!



ПЛАТФОРМА NEUROMATION

Предназначение платформы Neuromation - службы и данные для ИИ.



Обзор элементов платформы

Платформа Neuromation объединит технологические и бизнес-элементы для эффективной тренировки крупномасштабных систем глубокого обучения на искусственных данных.

Пользовательский интерфейс платформы разделен на следующие категории:



Модуль наборов искусственных данных:

ПРОЦЕССЫ:

- Создание генератора данных
- Заказ на создание набора данных при помощи генератора данных
- Запрос на маркировку данных

БИБЛИОТЕКИ:

- Репозиторий наборов данных для глубокого обучения
- Репозиторий генераторов данных
- Наборы данных (торговая площадка)



Модуль машинного обучения

ПРОЦЕССЫ:

- Определение модели глубокого обучения
- Импорт модели (например, клонирование кода TensorFlow)
- Заказ обучения по выбранному набору данных
- Запрос на пользовательскую модель на торговой площадке

БИБЛИОТЕКИ:

- Репозиторий моделей глубокого обучения (торговая площадка)



Пользовательский модуль

ПРОЦЕССЫ:

- Покупка токенов
- Регистрация в качестве потребителя или поставщика вычислительной мощности
- Регистрация в качестве потребителя или поставщика услуг
- Загрузка и установка программного обеспечения Neuromation Node

БИБЛИОТЕКИ:

- Данные пользователей
- Модели пользователей

Обзор элементов платформы

Серверная технология платформы разделена на следующие компоненты:



Узел Neuromation Node:

- Анализ узлов (минимальный уровень цен за обработку – минимальная сумма в нейротокенах (Neuro Tokens (NTK), принимаемая в качестве оплаты за единицу вычислений. Предельная производительность узла на основе пропускной способности, вычислительной мощности, емкости хранения)
- Пакет для распределенной генерации данных
- Пакет для распределенного обучения
- Промежуточное программное обеспечение для синхронизации узлов



Торговый (Рыночный) Модуль

- Обеспечение эффективного сопоставления заказов на покупку и продажу наборов данных, моделей, услуг маркировки
- Обеспечение ликвидности в системе, чтобы цены на активы способствовали масштабированию системы

Обзор биржевой торговой площадки ИИ

Во всех обменных операциях на платформе Neuromation используются наши токены.

Доступные услуги:

Покупка
данных

Классификация
данных

Генерация
данных

Покупка
модели

Тренировка
модели

Формирование цен в рамках платформы:

Цена каждой услуги будет определяться совокупным состоянием узлов Neuromation (цена за единицу вычислений). Для каждого узла будет установлена минимальная цена в токенах. Минимальный уровень цен можно также динамически регулировать с помощью алгоритма, который обеспечивает максимальную сумму токенов, зарабатываемых узлом. Платформа Neuromation будет определять ресурсы, необходимые для каждой запрошенной задачи, и выбирать наиболее эффективный пул узлов (минимизируя цену для заказчика). Платформа будет устанавливать равновесие предложения со стороны узлов, ищущих возможности на рынке, и спроса со стороны заказчиков, заинтересованных в оптимальной цене.

Покупка токенов клиентами:

Для совершения транзакции на платформе Neuromation клиенту необходимо будет купить токены. Чтобы упростить механизм покупки, Neuromation обеспечит клиентский портал, на котором покупка токенов будет производиться одним щелчком мыши.

Обратный выкуп токенов:

Для обеспечения ликвидности и регулирования рынка Neuromation будет периодически выкупать токены с рынка.





ЭКОНОМИЯ NEUROMATION

Майнеры будут получать больший доход за майнинг интеллектуальных данных для Neuromation, чем за майнинг криптовалюты

ОБМЕН КРИПТОВАЛЮТЫ



КЛИЕНТ



ETHER / BITCOIN

Neuro Tokens

УЗЛЫ NEUROMATION (управляемые майнерами Neuro Token'a)



УСЛУГИ

ГЕНЕРАЦИЯ ДАННЫХ
РАЗМЕТКА ДАННЫХ
ПОКУПКА ДАННЫХ
ОБУЧЕНИЕ МОДЕЛЕЙ
ПОКУПКА МОДЕЛЕЙ

Neuro Tokens

Neuro Tokens



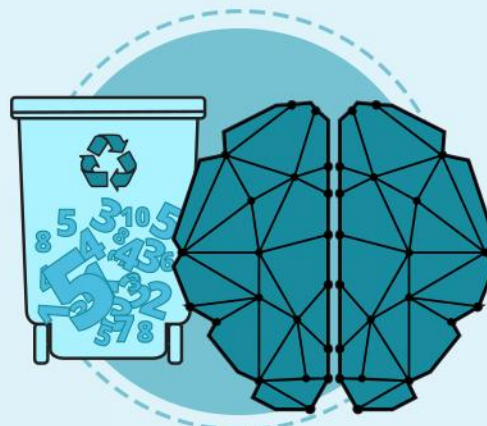
ETHER / BITCOIN

ОБМЕН КРИПТОВАЛЮТЫ



Neuro Tokens

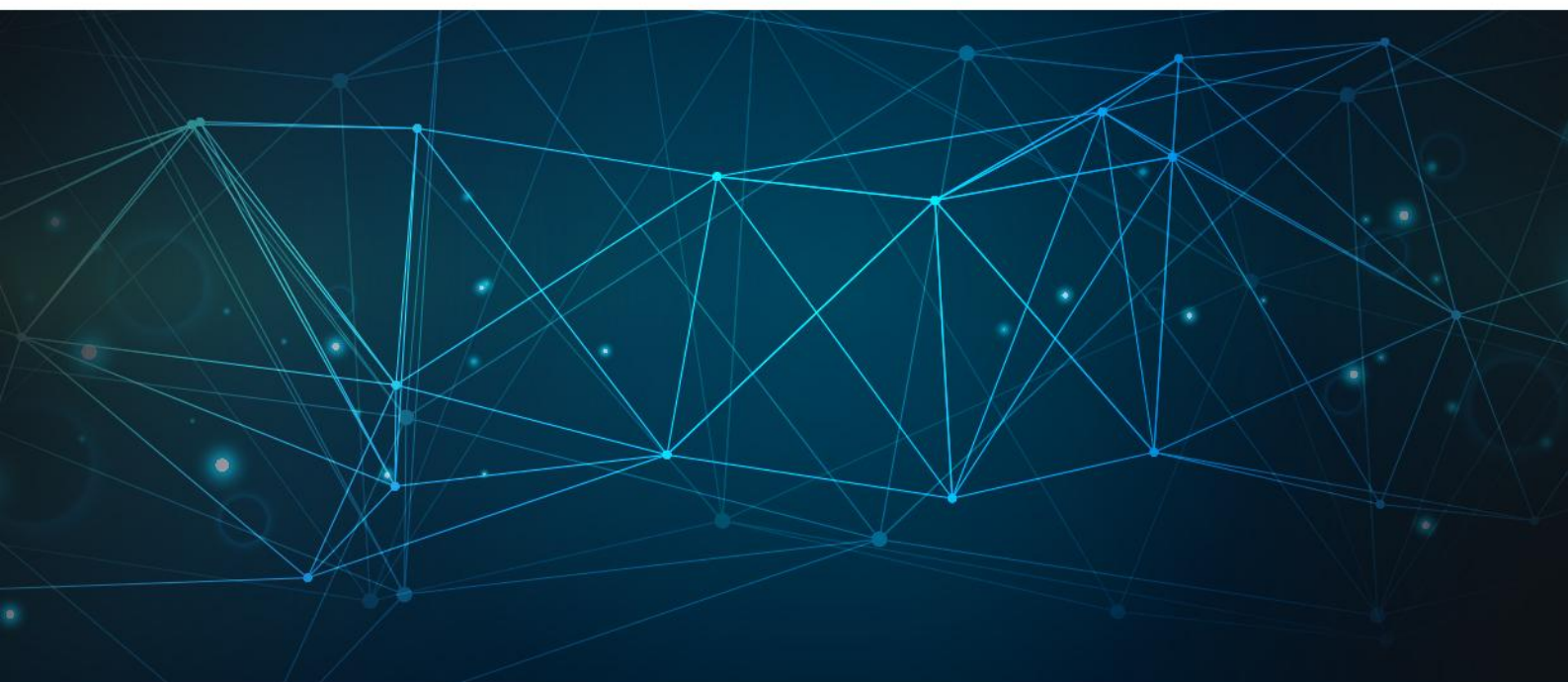
МАЙНИНГ ИНТЕЛЛЕКТУАЛЬНЫХ
ДААННЫХ ВМЕСТО МАЙНИНГА
КРИПТОВАЛЮТЫ



Нейронные сети и глубокое обучение

Глубокое обучение использует искусственные нейронные сети чрезвычайно большого размера; для обучения таких сетей часто требуются огромные наборы данных с высокоточной маркировкой. Сбор больших наборов данных изображений, текста или звука прост, но описание и аннотирование данных, помогающее сделать их пригодными для использования, традиционно оставалось сложным и дорогостоящим. Для решения проблемы сбора и маркировки наборов данных несколько лет назад был применен краудсорсинг, предполагающий привлечение большого количества людей для исправления ошибок и повышения точности. Это решение оказалось медленным, дорогостоящим, и результаты искажались за счет человеческого фактора. Мы предлагаем решение, в котором точность обеспечивается его конструкцией: синтезирование больших наборов данных с идеально точной маркировкой. Преимущества искусственных данных многообразны. Их можно быстро синтезировать и преобразовывать, они совершенно точны, адаптированы для текущей задачи, и их можно модифицировать для улучшения модели и самого процесса обучения.

В какой-то степени в отрасли уже используются искусственные данные, но это использование ограничено. В настоящее время нет общего набора инструментов, который мог бы помочь использовать искусственные данные в больших масштабах. **Платформа Neuromation изменит это.**



Нейронные сети и глубокое обучение

Десять лет назад машинное обучение пережило революцию. Хотя нейронные сети и являются одной из старейших парадигм в области искусственного интеллекта, исследователи и практики не смогли надежно и эффективно тренировать крупномасштабные нейронные сети до 2009 года. Благодаря новым идеям, выдвинутым в то время группами Джеффри Хинтона (Geoffrey Hinton), Йошуа Бенджио (Yoshua Bengio) и Яна ЛеКуна (Yann LeCun), нейронные архитектуры быстро превзошли самые передовые достижения в области обработки изображений, распознавания речи и обработки естественного языка, что привело к прорывным результатам в машинном обучении — искусственные нейронные сети перешли на **глубокое обучение**. Сегодня глубокое обучение — это *лингва-франка* в машинном обучении, оно охватывает самые различные сферы - от биометрии и беспилотных автомобилей до игры в Go и синтеза речи.

Основным строительным блоком нейронной сети является перцептрон - простая математическая модель нейрона, впервые предложенная Фрэнком Розенблаттом (Frank Rosenblatt) в 1957 году. Перцептрон вычисляет взвешенную сумму сигналов на своих входах (дендритах в биологических нейронах), веса которых представляют собой параметры модели, подлежащие изучению, за которыми следует нелинейная активационная функция. Сама нейронная сеть представляет собой большой набор перцептронов, которые могут быть совместно оптимизированы при помощи разных вариантов градиентного спуска. Градиенты для этого вычисляются методом обратного распространения, впервые предложенным для нейросетей в 1960-х годах и доведенным до полной общности в 1970-х годах.

Нейронные сети и глубокое обучение

Структура реализации нейронных сетей позволяет распараллеливать как обучение, так и применение таких моделей. Параллельная обработка данных в слоях сетей позволила значительно ускорить вычисления. Особенно важным стал переход к использованию графических процессоров, что привело к прорыву в высокопроизводительных вычислениях. В настоящее время практически все современные нейронные сети обучаются и работают на графических процессорах, а для следующего скачка эффективности может потребоваться специализированное оборудование – например, разработанный в Google тензорный процессор (TPU) или другие аппаратные платформы.

Известно, что даже самые простые нейронные архитектуры (например, двухслойная сеть прямого распространения) являются универсальными аппроксиматорами – они теоретически способны приблизить любую заданную математическую функцию. Однако чтобы иметь возможность аппроксимировать сложную функцию, сеть должна иметь достаточные выразительные возможности (емкость), и глубокие сети позволяют достичь высокой емкости более эффективно.

При попытке эффективно обучать глубокие нейронные сети большой емкости сразу же возникает проблема: *Где взять большие объемы хорошо маркированных данных?* и *Насколько точны эти метки?*

Эта проблема решается платформой Neuromation.

Искусственные данные

Одной из наиболее важных проблем современного глубокого обучения является получение *данных*. Обычные нейронные архитектуры, как правило, рассчитаны на *обучение с учителем*, тренировку для решения конкретной проблемы классификации или регрессии на размеченном наборе данных. В рамках глубокого обучения существуют методы *обучения без учителя* (например, стеки ограниченных машин Больцмана или порождающие состязательные сети, Generative Adversarial Networks) и подходы к преобразованию обучения без учителя в задачу обучения с учителем (при помощи автокодеров). Тем не менее, для получения практически значимых результатов требуется тщательная настройка с размеченными наборами данных. В современных исследованиях глубокого обучения много внимания уделяется *переносу обучения (transfer learning)*, повторному использованию обученных моделей для других задач на аналогичных данных, *адаптации предметной области (domain adaptation)*, повторному использованию обученных моделей для аналогичных задач на наборах данных с другими свойствами, и *однократному обучению (one-shot learning)*, использованию ограниченного числа примеров для обучения новому классу сети, уже обученной другим классам. Но эти усилия все еще далеки от мейнстрима.

Что же до настоящих данных, то хотя получить практически неограниченный набор неразмеченных данных (например, случайные фотографии из Интернета или записи человеческой речи) часто бывает легко, размечать такие данные вручную очень сложно и дорого. В течение последних нескольких лет краудсорсинг был практически единственным способом создания и разметки наборов данных, когда большое число людей исправляло ошибки и улучшало точность. Этот процесс оказался медленным и дорогостоящим, и результаты искажались человеческим фактором.

Средний розничный набор данных насчитывает свыше 150 000 позиций. Для глубокой нейросети требуются тысячи маркированных фотографий по каждой позиции. Чтобы сделать это вручную, потребуются годы.

Более того, некоторые данные (например, значения третьей координаты, глубины на фотографиях) практически невозможно разметить человеческими усилиями (мы плохо умеем оценивать абсолютную глубину), и для получения достоверных данных (точных меток) требуются сложные методы и инструменты.

Искусственные данные

Мы предлагаем решение, обеспечивающее идеальное качество меток благодаря своей конструкции: синтезирование больших наборов искусственных данных с идеально точной разметкой. У искусственных данных много преимуществ:

- как только среда готова, формирование данных в необходимых объемах происходит быстро и без значительных затрат;
- данные совершенно точны и адаптированы в соответствии с поставленной задачей, с разметкой, которую невозможно получить вручную (например, с картой освещенности окружающей среды или значениями глубины);
- их можно изменять для улучшения самих моделей и их обучения, создавая тем самым непрерывную обратную связь между моделью и искусственной средой.

В нашем решении для обучения используются искусственные данные, а небольшая тестовая выборка состоит из реальных данных с ручной разметкой. В сущности, это сводит использование искусственных данных к проблеме *переноса обучения*: нам нужно повторно использовать модели, обученные на одном наборе (искусственных) данных, и применить их к другому типу набора данных (реальным изображениям). Однако наш подход имеет несколько важных преимуществ, которые в этом случае значительно упрощают перенос обучения. Во-первых, набор искусственных данных обучения не привносится как часть проблемы, а генерируется нами самими, и мы можем и стремимся сделать его сопоставимым с реальными данными.

Данные — это новая нефть. Получение новых наборов данных с ручной маркировкой сравнимо с поиском новых месторождений нефти. Искусственные данные похожи на изобретение дешевой и повсеместно распространенной синтетической нефти.

Во-вторых, что еще более важно, вышеупомянутая петля обратной связи между моделью и набором данных значительно упрощает перенос обучения: мы можем настраивать не только модель, но и тренировочный набор данных — роскошь, редко доступная в практике машинного обучения.

Искусственные данные

В настоящее время при применении этого подхода мы используем самые передовые модели обнаружения объектов и сегментации изображений, чтобы распознавать предметы на полках продуктового магазина / супермаркета. Среди этих объектов - такие, как бутылки безалкогольных напитков (Pepsi/Mirinda), коробки с соками, пакеты картофельных чипсов и т. д. Это конкретное приложение особенно интересно с точки зрения тренировки моделей на искусственных данных. Сами объекты относительно легко визуализировать: бутылка Pepsi представляет собой намного более простую трехмерную модель, чем человеческое лицо.

Данные с ручной маркировкой в достаточных объемах очень сложно получить по следующим причинам:

- крупнейшие пищевые компании, такие как Pepsico, могут производить тысячи разных наименований продукции, и их каталоги часто меняются;
- задача состоит в том, чтобы иметь возможность распознавать эти объекты на полках супермаркетов, с сотнями предметов на каждой, по фотографиям; ручная маркировка даже одной такой фотографии – это трудоемкий процесс, а их нужны миллионы;
- для высококачественного обучения также требуются большие наборы нулевых данных, маркированных объектов, схожих с теми, что являются предметом обучения, но не являющихся таковыми; обычно они 10-кратно превосходят тренировочный набор.



Искусственные данные

Мы произвели тренировку несколько различных моделей *распознавания объектов* (SSD, YoLo, Fast(er) R-CNN), *сегментации объектов* (SegNet, FCN для сегментации) и в настоящее время работаем над реализацией Mask R-CNN для *пообъектной сегментации (instance segmentation)*. Все эти модели содержат сверточную сеть, которая выполняет основную часть работы по категоризации изображения, и отличаются дополнительными слоями сверх этих функций, которые обрабатывают собственно семантику и производят метки классов, предположительное разграничение объектов, сегментацию и другие исходные параметры.

Мы уже получили отличные результаты распознавания объектов как на искусственных тестовых наборах (что и следовало ожидать), так и на тестовых наборах реальных фотографий (что обнадеживает и служит доказательством правильности концепции).

В дальнейшей работе мы планируем улучшить конвейеры генерации искусственных данных с помощью алгоритмов машинного обучения, что приведет к подлинно активной структуре обучения: искусственные данные используются для улучшения модели, которая проверяется на реальных данных, а, в свою очередь, обученная модель улучшает конвейер генерации искусственных данных для дальнейшего совершенствования обучения.



Распределенная обработка данных и экономика Neuromation

Недостаток вычислительной мощности, необходимой для масштабной генерации искусственных наборов данных, является ключевым препятствием для широкого внедрения. У платформы Neuromation есть экономически обоснованное предложение для майнеров криптовалют, которое поможет привлечь их вычислительные ресурсы для решения этой проблемы.

Предыстория: На майнеров криптовалют оказывает давление все уменьшающаяся эффективность их вычислений proof-of-work, необходимых при майнинге биткойнов, эфиров или других криптовалют. Растущая конкуренция, консолидация майнинг-пулов и растущий спрос на вычислительные мощности и ресурсы в скором времени не позволят многим мелким майнерам вести экономически оправданную деятельность.

Экономическое стимулирование майнеров со стороны Neuromation: Мы хотим, чтобы у майнеров были варианты. Помимо своего существующего программного обеспечения для майнинга, они могут загрузить Neuromation Computation Node. Когда предлагается выполнить задачу Neuromation, любой узел может участвовать в торгах. Если узел выигрывает конкурс, он переключает свои вычислительные мощности с майнинга эфиров или биткойнов на задачу блокчейн-платформы Neuromation. Узел генерирует искусственные данные или занимается тренировкой модели глубокого обучения. В качестве награды майнер получает наши токены: Криптовалюта Neuromation. Как только задание выполнено, работа Neuromation Node прекращается и продолжается майнинг криптовалюты. Первоначально наши жетоны будут расширением системы Эфириум, но позже, когда платформа достигнет экономической зрелости, мы перейдем на свой собственный блокчейн. Наши предварительные оценки показывают, что на токенах Neuromation майнеры будут зарабатывать в 3–5 раз больше, чем на майнинге криптовалют. Майнер не будет занят задачами Neuromation 100% времени, поэтому это будет повышением эффективности использования существующего оборудования. Ниже приведена разница в реальной доходности аналогично сконфигурированных майнинг-ферм. На одной работает алгоритм майнинга криптовалюты, а на другой запущена задача Deep Learning / Data Rendering.

Тип вычислений	Доход за ед. времени
Майнинг биткойнов	\$ 7 - 8 Долл. США/День
Amazon Deep Learning	\$ 3 - 4 Долл. США/Час

Модели машинного обучения

- В широком доступе находится множество моделей: программы компьютерного зрения **VGG, GoogLeNet** и **ResNet**, программы генерирования речи типа **Wavenet** и т. д.
- Подобное ПО интегрируется в различные фреймворки и библиотеки, например, **TensorFlow, Caffe, Torch, pyTorch, Theano, Keras**, а для связи разнотипных фреймворков применяют конвертеры.
- Платформа Neuromation будет поддерживать все основные фреймворки и библиотеки глубокого обучения и автоматического дифференцирования, а также основные предварительно обученные модели.
- Результаты обучения будут доступны для скачивания.
- К моделям применяется сжатие для уменьшения времени отклика или при использовании в мобильных устройствах. Методика **сжатия моделей** включает:
 - удаление избыточной информации в слоях после завершения обучения (крупной) модели. Коэффициент сжатия сети достигает 2-3, что особенно актуально в очень глубоких сетях компьютерного зрения, основанных на остаточных связях. Метод называется «**очистка модели**»;
 - некоторые сети, известные также как «**мобильные сети**», изначально разрабатываются небольшими (напр., сеть распознавания образов SqueezeNet).
- Заявки на модели и архитектуры специального исполнения. Клиент может разместить запрос на разработку модели с особыми характеристиками, если существующие модели и инструменты не подходят к конкретному приложению. Этот запрос размещается на **бирже Neuromation** и будет обработан специалистами Neuromation или другими сертифицированными поставщиками.

Лаборатории Neuromation

Наша стратегия заключается не в том, чтобы развить нашу платформу в изоляции, а в том, чтобы работать с партнерами в избранных отраслях и попытаться органично воплотить нашу концепцию в жизнь. Мы готовимся открыть Neuromation Labs (лаборатории Neuromation), которые будут заниматься выработкой искусственных данных и тренировать модели глубокого обучения в реальных приложениях. Каждая "лаборатория" будет посвящена исследованию конкретной проблемы в сотрудничестве с лидером в своей категории. По мере того как наша платформа будет развиваться, мы будем перемещать в Labs часть генерации и обучения, что позволит нам органично тестировать части нашей концепции в реальных сценариях.

Ключ к нашему успеху – в сотрудничестве с бизнесом

На начальном этапе лаборатории будут заполнять рынок платформы Neuromation генераторами и наборами данных. Мы также будем поощрять наших партнеров по Labs к предоставлению дополнительных услуг через платформу и созданию, таким образом, первоначального рынка. Мы уже запустили две лаборатории и продолжаем работу над следующими. Работа с примерами из реального мира, когда отраслевые партнеры формируют совместно с нами "лабораторию", позволяет нам обобщать близкие подходы из различных отраслей и интегрировать их в нашу платформу, делая ее более мощной.

Пример модели Neuromation — классификация товаров на полках супермаркетов



"Электричество преобразило почти все 100 лет назад, сегодня, как я думаю, ИИ — это новое электричество", — [слегка перефразировано] Эндрю Нг (Andrew Ng)

Автоматизация розничной торговли

Neuromation, в сотрудничестве с ведущими брендами в отрасли, занимается решением проблемы поиска потребительских товаров на полках магазинов. Используя искусственные данные по каждой товарной позиции, мы можем создавать большие идеально маркированные наборы данных для сотен тысяч SKU в розничной торговле. Модели глубокого обучения для компьютерного зрения, прошедшие тренировку на этих больших наборах данных, способны эффективно анализировать и классифицировать доступность полок, процент заполнения, точность раскладки и другие показатели.

Автоматизация розничной торговли(live)

Наша лаборатория технологий розничной торговли первой выходит в реальную жизнь в сотрудничестве с избранными партнерами на рынках Европы, Ближнего Востока, Северной и Южной Америки. В идеале каждая полка будет в реальном времени передавать поток аналитических данных нашим партнерам-ритейлерам. Используя нашу концепцию искусственных данных, мы можем быстро и точно производить перетренировку аналитических моделей, делая ИИ гибким и полезным инструментом для применения в различных сферах розничной торговли.

Пример искусственных данных — полученные в результате процедурной генерации трехмерные полки магазинов:



Ожидается, что в ближайшие 15 лет будет продолжаться экспоненциальный рост на готовом принять инвестиции рынке роботизации складского хозяйства и логистики.

Автоматизация розничной торговли (live)

Радует, что всего за месяц работы над проблемой распознавания товарных позиций мы достигли точности свыше 95% — результата, на который у других ушли годы усилий и миллионы долларов. Примечательно, что модель хорошо работает с реальными данными всего лишь после тренировки на искусственных наборах данных. Этот прорыв доказывает жизнеспособность и эффективность нашего подхода.

С появлением платформы Neuromation откроется, наконец, доступ к облегчению широкомасштабной тренировки ИИ.

(Прорыв) Модель, обученная на искусственных данных, распознает реальные данные.



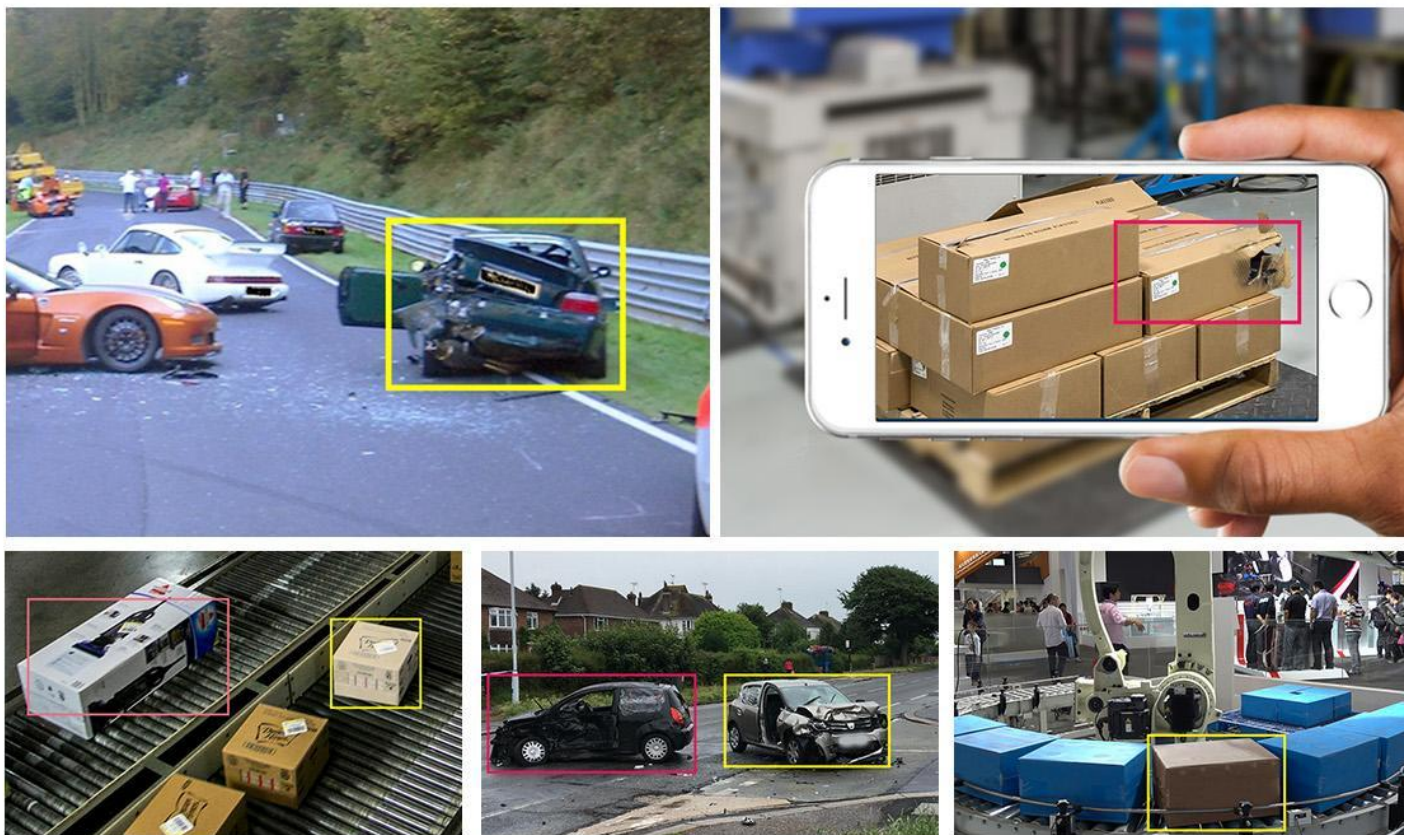
Потенциальный спрос только на распознавание изображений в индустрии розничной торговли огромен.

По словам Гартнера (Gartner), к 2020 году 85% клиентских взаимодействий в розничной торговле будет производиться искусственным интеллектом.

По данным исследования 72 крупнейших розничных торговцев и поставщиков, проведенного ECR, этот объем составляет более 40 миллиардов изображений в год. В дальнейшем мы планируем создать наборы данных, которые имитируют взаимодействие человека с полкой. Создав полную имитацию розничного магазина, мы сможем отслеживать поток клиентов и их намерения, и для них у нас найдется множество разнообразных приложений.

Промышленная автоматизация

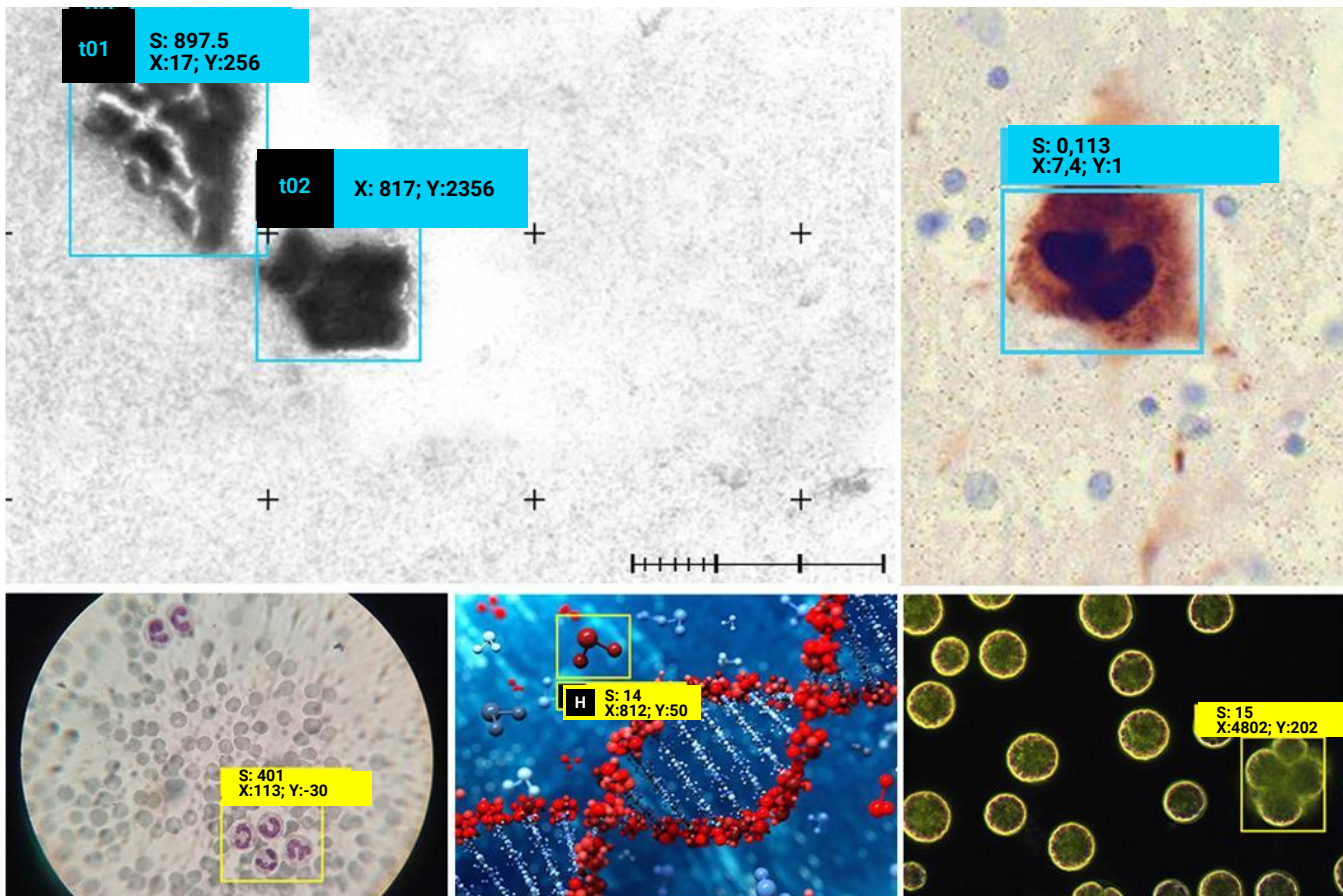
Типичной проблемой в промышленной среде является высокая динамичность выполняемых задач. Модели должны быть достаточно гибкими, чтобы работать с динамическими данными в реальном времени, обеспечивая высокую точность на выходе. Эта обработка выполняется "на лету" (иногда в буквальном смысле – некоторые задачи выполняются беспилотными летательными аппаратами). Наборы данных в реальном времени не могут быть построены в динамической среде достаточно быстро, со скоростью, необходимой для своевременной обработки и вывода. Решение этой проблемы – это **платформа Neuromation**, которая позволяет процедурно создавать виртуальные среды, в которых модели будут проходить тренировку. После тренировки в этих виртуальных песочницах модели можно быстро откалибровать на реальных данных. Гибкая проверка гипотез становится возможной.



Neuromation планирует открыть лабораторию Enterprise Automation (промышленной автоматизации), где концепция искусственных данных поможет внедрять новые решения в производство, цепочку поставок, финансовые услуги и сельскохозяйственные отрасли, и это лишь некоторые примеры.

Фармакология / Медицина / Биотехнологии

Для концепции Neuromation есть множество интересных объектов приложения в сфере фармацевтики и биотехнологий. В общем, любая задача, где построение проще распознавания (результат легко рассчитать по входным данным, а обратная операция сложна, подобно хеш-функциям в информатике), хорошо решается с помощью нашей концепции искусственных данных. Мы полагаем, что масса задач в ходе построения гипотез при разработке лекарственных средств и в динамике биологических систем может быть решена с использованием моделей глубокого обучения. Мы запускаем эту лабораторию в сотрудничестве с ведущим поставщиком решений для наблюдения за младенцами.



Мы будем совместно разрабатывать смарт-камеру, которая будет отправлять предупреждения, если ребенок перевернулся, изменил характер дыхания, проснулся и т. д.

Другие лаборатории Neuromation

Neuromation планирует открыть и другие лаборатории в партнерстве с промышленными предприятиями. Наша стратегия состоит в том, чтобы иметь партнера в каждой области.



Максим Прасолов
Исполнительный директор

Максим, топ-менеджер и продюсер, в период с 2001 по 2017 год продюсировал более 50 мультипликационных рекламных роликов, 3D-фильмов, коммерческих и промышленных приложений, компьютерных игр. Он работал с международными торговыми и промышленными брендами, такими как Unilever, Yukos, TPE, Metro Cash & Carry, Severstal Group, Ferrexpo, коммерческими банками, инвестиционными и страховыми компаниями. Макс был членом команды, осуществившей первоначальное публичное размещение (IPO) на Лондонской фондовой бирже акций Ferrexpo, крупнейшего производителя железной руды в Украине. С 2014 года он инвестирует в стартапы в области разработки беспилотных летательных аппаратов, ИИ, мультимедиа с использованием дополненной реальности (AR).



Федор Савченко
Технический директор

У Федора более чем 20-летний опыт руководства сложными проектами разработки программного обеспечения с акцентом на компьютерной графике, **3D-движках**, производстве CGI и виртуальной реальности (VR). Он имеет ученые степени в области математики и графического дизайна. Это уникальное сочетание позволяет ему создавать программное обеспечение, которое может быть одновременно и функциональным, и красивым. Он - новатор, и всегда находится на передовой последних тенденций.



Сергей Николенко
Директор по науке

Научный сотрудник Математического института им. В. А. Стеклова в области машинного обучения (глубокое обучение, байесовские методы, обработка естественного языка и др.) и анализа алгоритмов (сетевые алгоритмы, конкурентный анализ). Сергей является автором более чем 120 научных работ, нескольких книг, курсов "Машинное обучение", "Глубокое обучение" и других. У него обширный опыт участия в коммерческих проектах (Neuromation, SolidOpinion, Surfingbird, Deloitte Analytics Institute).



Константин Гольцев
Инвестор /
Председатель

Профессиональный предприниматель и ветеран индустрии онлайн-рекламы, Константин имеет более чем 20-летний опыт разработки программного обеспечения и продукции. Он был генеральным директором и основателем новаторской рекламной сети AdoTube, которая выросла с нуля до 200 сотрудников, занятых продажами на рынках 23 стран, с 13 офисами по всему миру. AdoTube была впоследствии продана Exponential Interactive. Он основатель и президент SolidOpinion - компании, которая революционизировала привлечение внимания онлайн-сообществ к контенту.



Денис Попов
Директор по
информационным
технологиям

У Дениса более чем 15-летний опыт разработки программного обеспечения и руководства крупными техническими проектами от их зарождения до завершения. Он был техническим директором, главным научным советником (CSA), вице-президентом по инженерным разработкам и консультантом в десятках успешных проектов и стартапов. Его вклад в работу состоит в опыте управления техническими процессами и познаниях в сферах многоуровневого параллелизма, хранения больших данных и анализа данных в реальном времени. Денис был техническим руководителем в компании Viewdle, которая занималась распознаванием лиц и была приобретена Google.



Кирилл Трусовский
Ведущий научный
сотрудник

Кирилл обладает более чем четырёхлетним опытом в области машинного обучения. Огромная часть его заслуг связана с созданием стартапов с нуля. У него глубокое понимание выбора и реализации полной архитектуры и крупномасштабных решений. Крепкий математический фундамент и страсть к исследованиям побудили его продолжить учебу, и теперь Кирилл является аспирантом, занимаясь научными исследованиями в области компьютерных наук.



Эстер Кац
Директор по
коммуникациям

Специалист по связям с общественностью с 20-летним опытом работы на основных рынках, специализируется на био- и финансовых технологиях. В недалеком прошлом - глава Jacob Peres Office, ведущего PR-агентства в Израиле. Давний приверженец криптовалют и энтузиаст технологии блокчейн. Живет и работает в Израиле.



Эван Кац
CRO

Эван много лет проработал в самых быстрорастущих медиа-стартапах Израиля. До прихода в Neuromation он стал одним основателей успешной медиа-компании, использующей ИИ. Он предприниматель с обширным опытом в сфере цифровых медиа, развития бизнеса и продаж. Эти качества сделали его незаменимым участником многих успешных стартапов. Он амбициозен, решителен, целеустремлен и привносит в любой проект свой энтузиазм, харизму и обаятельность, что привлекает к нему людей и помогает добиваться успеха.



Эндрю Рабинович
Консультант

Ученый с мировым именем в области исследования глубокого обучения и машинного зрения. Уже более 15 лет Эндрю занимается исследованиями в области машинного обучения с упором на машинное зрение. Он автор многочисленных патентов и научных публикаций. Основал стартап в области биотехнологии, который был успешно продан. Эндрю получил степень доктора философии (PhD) в области компьютерных наук в Калифорнийском Университете в Сан-Диего в 2008 году. Он много лет работал на ведущих позициях в области научно-исследовательских разработок в Google, Magic Leap, где в настоящее время является директором по глубокому обучению.



Дэвид Орбан
Консультант

Дэвид – учредитель и управляющий партнер международной инвестиционной фирмы «Network Society Ventures», специализирующейся на посевном финансировании инновационных компаний, работающих на стыке новейших технологий и децентрализованных сетей. Бизнесмен, тонкий стратег, гуру, автор, блогер, докладчик и авторитетный эксперт в области глобальной технологической среды. Предпринимательский талант Дэвида реализован в ряде компаний, успешно развивающихся уже более 20 лет. Дэвид является учредителем и членом правления международной научно-исследовательской организации «Network Society Research» (Лондон).



Юрий Кундин
Консультант

Юрий Кундин (Yuri Kundin), директор офиса KPMG US в Сан-Франциско, специализируется на консультационных услугах в области управления рисками и соответствия нормативно-правовым требованиям. У Юрия более чем 15-летний опыт консультирования по вопросам соблюдения нормативных требований различных агентств США, таких как SEC, OCC и FDIC, а также валютных управлений Сингапура и Гонконга и Центрального банка России. В последнее время Юрий специализируется на разработке структуры и методологии оценки рисков, соответствия и аттестации экосистем блокчейнов, криптовалют и ICO.



Дэйв Карлсон
Консультант

Дэйв Карлсон – исполнительный директор Giga-Watt, одного из крупнейших в Северной Америке поставщиков комплексных услуг и индивидуальных решений в сфере майнинга. Дэйв был одним из первопроходцев в сообществе майнеров, обладает большим опытом в области разработки программного обеспечения и является успешным. Прежде чем занять должность исполнительного директора Giga-Watt, Дэйв стал основателем и в течение 4 лет работал исполнительным директором в MegaBigPower, одном из крупнейших в мире независимых операторов мощностей майнинга

Источники дохода - платформа

Одна из целей платформы Neuromation – стать ключевым центром сервисов искусственного интеллекта для международного бизнеса. Ожидается, валовый объем транзакций на платформе составит \$71 млн. С каждой транзакции Neuromation получит от 5% до 15% комиссионных в зависимости от типа предоставленной через платформу услуги.

В 2018–2022 гг. использование платформы будет увеличиваться в 3-5 раз ежегодно, а через три года ежегодный доход от комиссионных средств превысит 100 млн.

ОЦЕНОЧНЫЙ АНАЛИЗ АКТИВНОСТИ ПЛАТФОРМЫ (*примечание)

Операции на платформе	Цена, ETH	Цена, Neurotoken(NTK)	Цена, USD	Время
Созд. генератора данных	100.00	100 000.00	20 000	4 недели
Генерирование данных (10 тыс.)	0.17	174.22	100	1 день
Обучение модели	0.01	87.11	50	0.5 дня
Маркировка данных (по 1 тыс.)	0.88	871.09	500	1 день
Продажа наборов данных	10.00	10.00	2 000	Н/Д
Продажа моделей	100.00	100.00	20 000	Н/Д

ОЖИДАЕМЫЙ ОБЪЕМ ОПЕРАЦИЙ НА ПЛАТФОРМЕ В 2018 ГОДУ (*примечание)

Операции на платформе	Ожидаемое кол-во операций, 2018	Количество, ETH	Количество, Neurotoken(NTK)	Количество, USD
Созд. генератора данных	1000	100 000.00	100 000 000.00	20 000 000
Генерирование данных (10 тыс.)	10 000	1 742.16	1 742 160.28	1 000 000
Обучение модели	100 000	1 451.80	8 710 801.39	5 000 000
Маркировка данных (по 1 тыс.)	10 000	8 710.80	8 710 801.39	5 000 000
Продажа наборов данных	10 000	100 000.00	100 000 000.00	20 000 000
Продажа моделей	1000	100 000.00	100 000 000.00	20 000 000
Всего	132 000	311 904.76	319 163 762.07	71 000 000

* Оценка данных приблизительная и может существенно отличаться от реальных показателей.

Источники дохода – лаборатории Neuromation

Лаборатории Neuromation сотрудничают с отраслевыми лидерами, продвигая решения, связанные с искусственным интеллектом. Часть прибыли Neuromation формируется за счет продажи API (SaaS) или данных промышленным партнерам.

Примеры:

- **Лаборатория розничных продаж**

Лаборатория заключила контракт с OSA HP (<http://osahp.com/en/>) и ECR (<http://ecr-all.org/?lang=en>) на поставку 170.000 массивов синтетических данных для использования в розничной торговле на территории Восточной Европы. Предусматривается создание моделей глубокого обучения, распознающих объекты на полках магазинов. Эти модели станут неотъемлемой частью аналитических инструментов для выявления доступности товара на полках. По итогам контракта прибыль составит 4,25 млн. евро за 1,5 года. С реализацией модели на мобильном устройстве можно ознакомиться здесь: <https://youtu.be/Q01KGwQKoSU>

- **Лаборатория приборов медицинского назначения**

В сотрудничестве с ведущим производителем устройств для слежения за грудными детьми MonBaby (<https://monbaby.com/>) планируется разработать смарт-камеру, которая расширит функционал продукта. Камера будет отслеживать движения ребенка и передавать родителям ценные данные. Neuromation будет получать долю прибыли от каждого проданного устройства. Доход от реализации контракта в течение нескольких лет превысит 2 млн. евро.

- **Другие лаборатории**

В начале 2018 года планируется запуск новых лабораторий, что создаст дополнительные источники прибыли для Neuromation.

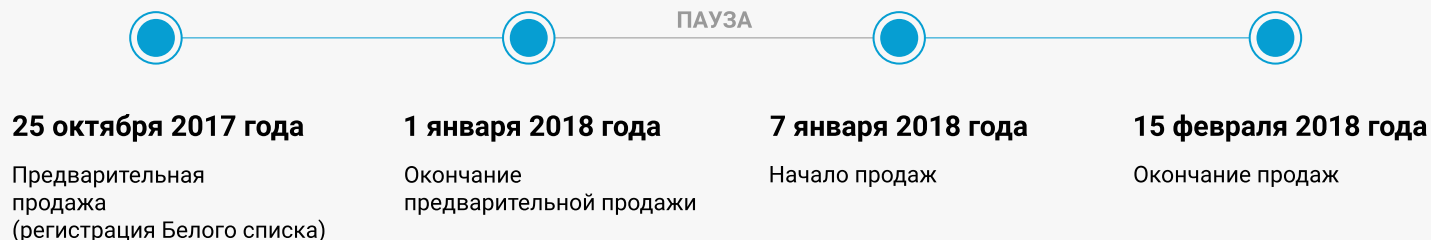
В целом прогноз следующий: партнерство лабораторий Neuromation с другими компаниями будет приносить около 30% совокупного дохода.



После 2018 года Neuromation становится признанным партнером бизнеса в развитии потенциала использования ИИ. Наши токены являются основным механизмом обмена, позволяющим генерировать искусственные данные, осуществлять распределенную тренировку моделей, маркировку данных и другие услуги в области ИИ.

Мы предполагаем стать глобальным пулом ресурсов в сфере искусственных данных, постоянно пополняющейся библиотекой, которая будет располагать наборами данных для любого возможного варианта использования. Поскольку мы связываем поставщиков услуг с платформой, Neuromation станет благоприятной экосистемой для любых сфер применения ИИ.

Временной график:



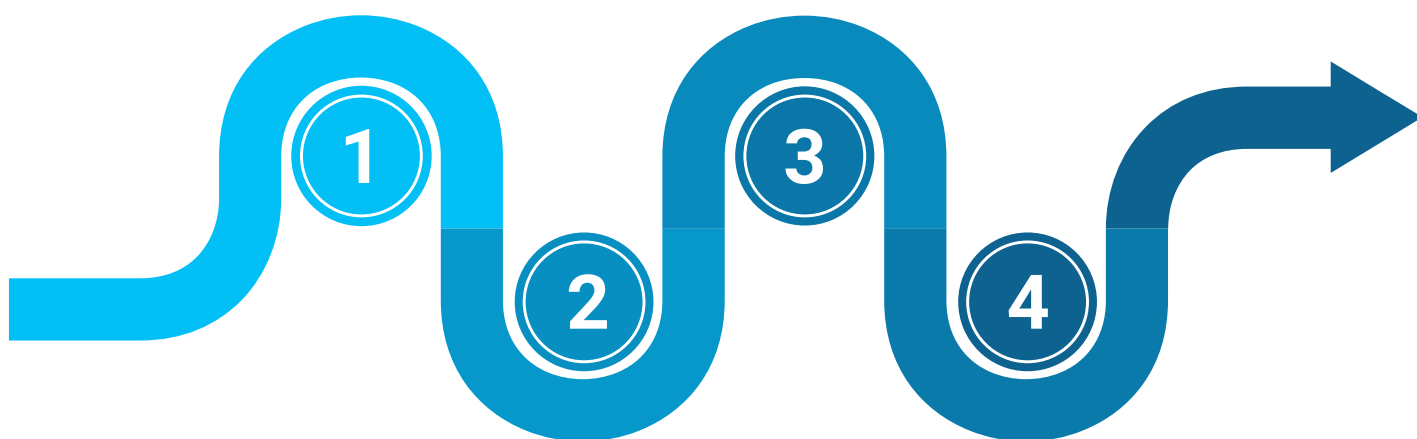
Цели:

Лаборатории Retail (розничной торговли) + Med-Devices (медицинских устройств).
Базовая платформа Neuromation.

2,000 ETH

Большая платформа Neuromation (полностью распределенная генерация данных и вычисления), пул из 10 000 GPU в рамках платформы.

10,000 ETH



5,000 ETH

Лаборатория Industrial Innovation (промышленных инноваций), начало работы кластера из 1000 GPU в рамках платформы. ☒

60,000 ETH

Фонд Neuromation, пул из 100 000 GPU. Всемирное распространение

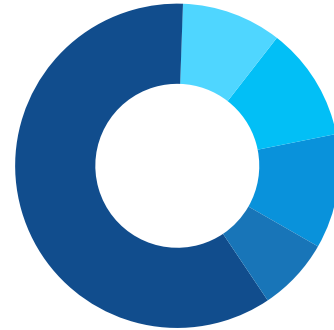
Юрисдикция: Эстония. Neuromation полностью соответствует законодательству Эстонии, регулирующему краудфандинг.

Количество эмитируемых токенов:

- Эмиссия токенов: 1 000 000 000 (дополнительных эмиссий не будет)
- Выпущено 60 000 000 токенов. Neuromation выведет из обращения токены, оставшиеся нераспроданными.

Вывод из обращения: 30% 2018, 20% 2019, 10% 2020 (Neuromation выведет из обращения 50% токенов в течение следующих трёх лет и сократит запас доступных токенов)

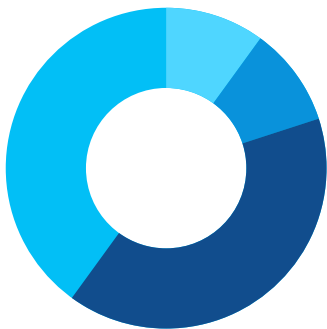
- Цена за 1 токен = 0,001 ETH



Распределение токенов:

- 60% в продаже токенов
- 12% команде
- 12% партнерам
- 10% - резерв ликвидности
- 6% исследователям

Распределение доходов от ICO:



- Не менее 40% — на развитие платформы
- До 40% — резерв ликвидности (предоплата за серверные мощности)
- 10% — на PR и маркетинг услуг Neuromation
- 10% — вознаграждение партнерам / консультантам / ранним инвесторам

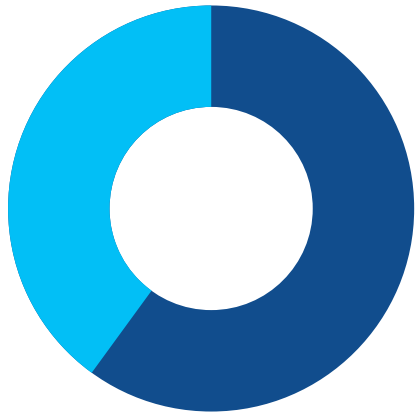
Бонусная программа ICO:

СРОКИ ВЫПЛАТЫ БОНУСА:

- Предпродажа (зарегистрируйтесь, чтобы узнать процент бонуса)
- Первая неделя: + 15% в токенах
- Вторая неделя: + 10% в токенах
- Третья неделя: + 5% в токенах

СУММА БОНУСА (НТК):

- 30 эфиров - 1%
- 50 эфиров - 2.5%
- 100 эфиров - 5%
- 300 эфиров - 8%
- 500 эфиров - 10%
- 1000 эфиров - 15%
- 2000 эфиров - 20%
- 3000 эфиров - 30%+



	Эмитировано токенов		Доступно на ICO
	100 000 000		60 000 000 (*)
	100 000 ETH		60 000 ETH

* Во период продажи токенов все непроданные токены в доступном пуле будут сожжены

Номинальная цена токена:

$$1 \text{ ⚡} = 0,001 \text{ ⬠}$$

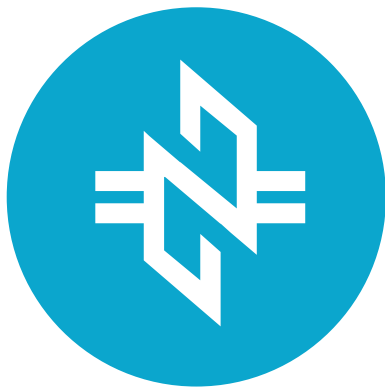
ОЦЕНОЧНЫЙ АНАЛИЗ АКТИВНОСТИ ПЛАТФОРМЫ (*примечание)

Операции на платформе	Цена, ETH	Цена, Neurotoken (NTK)	Цена, USD	Время
Созд. генератора данных	100.00	100 000.00	20 000	4 недели
Генерирование данных (10 тыс.)	0.17	174.22	100	1 день
Обучение модели	0.01	87.11	50	0.5 дня
Маркировка данных (по 1 тыс.)	0.88	871.09	500	1 день
Продажа наборов данных	10.00	10.00	2 000	Н/Д
Продажа моделей	100.00	100.00	20 000	Н/Д

ОЖИДАЕМЫЙ ОБЪЕМ ОПЕРАЦИЙ НА ПЛАТФОРМЕ В 2018 ГОДУ (*примечание)

Операции на платформе	Ожидаемое кол-во операций, 2018	Количество, ETH	Количество, Neurotoken(NTK)	Количество, USD
Созд. генератора данных	1000	100 000.00	100 000 000.00	20 000 000
Генерирование данных (10 тыс.)	10 000	1 742.16	1 742 160.28	1 000 000
Обучение модели	100 000	1 451.80	1 451 800.80	5 000 000
Маркировка данных (по 1 тыс.)	10 000	8 710.80	8 710 801.39	5 000 000
Продажа наборов данных	10 000	100 000.00	100 000 000.00	20 000 000
Продажа моделей	1000	100 000.00	100 000 000.00	20 000 000
Всего	132 000	311 904.76	311 904 760.00	71 000 000

* Это только оценка. Фактическая производительность может существенно отличаться.



2018 Спрос на Neurotoken (NTK)	2019 Спрос на Neurotoken(NTK) (2018 x5)	2020 Спрос на Neurotoken (NTK) (2019 x3)
319 163 763.07	1 595 818 815.33	4 787 456 445.99

Сожжено токенов (% от общего объема в обращении)

30%	20%	10%
-----	-----	-----

Всего доступно

70 000 000.00	56 000 000.00	50 400 000.00
---------------	---------------	---------------

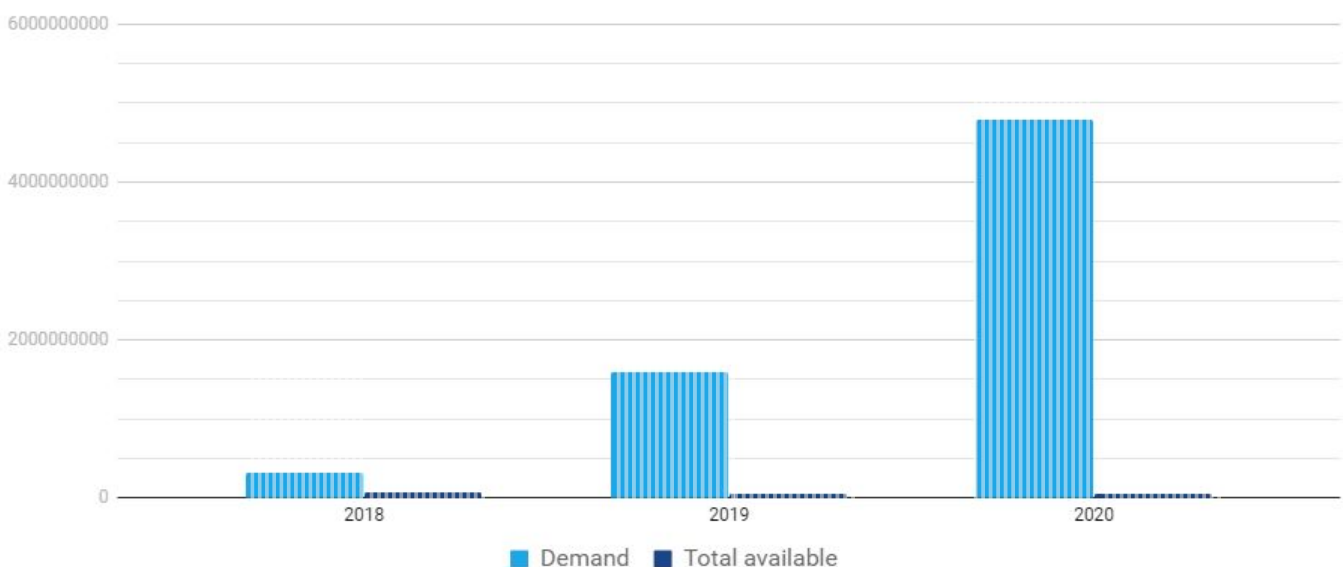
* Предполагаемое количество токенов в продаже - 60 000 000

Спрос на токены превышает предложение в 4 раза

Спрос на токены превышает предложение в 30 раз

Спрос на токены превышает предложение в 90 раз

Neuro Token demand



Согласно нашей политике сожжения токенов, спрос на Neurotoken на платформе Neuromation к концу 2020 года **возрастет в 90 раз.**