**ГЛОССАРИЙ**

### **ОБЩИЕ ТЕРМИНЫ**

**API** – это программный интерфейс, который задаёт описание способов взаимодействия одного программного модуля с другим;

**Baseline** – это начальное решение, которое служит опорой для дальнейшей работы.

**Big Data** (Большие данные) –это структурированные или неструктурированные массивы данных большого объема;

**Data** (Данные) - это совокупность сведений, зафиксированных на определённом носителе в форме, пригодной для постоянного хранения, передачи и обработки;

**Data Science** (Наука о данных) - это комбинация нескольких дисциплин, которая использует статистику, анализ данных и машинное обучение для анализа данных и извлечения из них знаний и полезных данных;

**Information** (Информация) – это данные имеющие ценность для субъекта;

**Pet-project** – это внерабочая деятельность, которая направленная на прокачку профессиональных способностей разработчика в свободное время по собственной инициативе.

**Web-service** – это программа в интернете, которая оказывает услугу или отвечает на определенное требование пользователя.

**Неуспеваемость** - это комплексная, суммарная, итоговая неподготовленность студента для дальнейшей учебы.

### 

### **МАТЕМАТИЧЕСКИЕ ТЕРМИНЫ**

**Coverage** (Генеральная совокупность) –это совокупность всех возможных объектов, которые подлежат изучению в пределах объекта исследования;

**Median** (Медиана) – это значение признака, которое делит упорядоченное множество данных пополам;

**Sampling** (Выборочная совокупность) –этомножество объектов определённым способом выбранных из генеральной совокупности для использования в исследовании;

**Гистограмма частот** – это столбчатая диаграмма отображающая зависимость частоты встречаемости переменной от её значения;

**Мода** – это значение признака, которое встречается чаще всего;

**Репрезентативность выборки** – это свойство отражающие сопоставимость выводов по выборке выводам генеральной совокупности по свойствам;

**Форма распределения** – это кривая, огибающая столбцы частотной гистограммы.

### **MACHINE LEARNING ТЕРМИНЫ**

**Machine Learning** (Машинное обучение) **–** этонаука создания алгоритмов (моделей), которые улучшаются благодаря опыту;

**Training set** (Обучающая выборка) – это обработанная и структурированная информация в виде таблицы;

**Model** (Модель) – это математическое представление системы или процесса;

**Метод fit-predict** – это использование модели без понимания её математического аппарата;

**Метрика качества** – это способ оценки производительности модели на основе её способности предсказывать верные ответы для тестовой выборки.

### **DEEP LEARNING ТЕРМИНЫ**

**Deep Learning** (Глубокое обучение) –этонаука создания алгоритмов (моделей), которые используют данные как опыт.