



СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

Александр Калиниченко

ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

Модуль 2. Методы искусственного интеллекта

Лекция 13. Рекуррентные нейронные сети

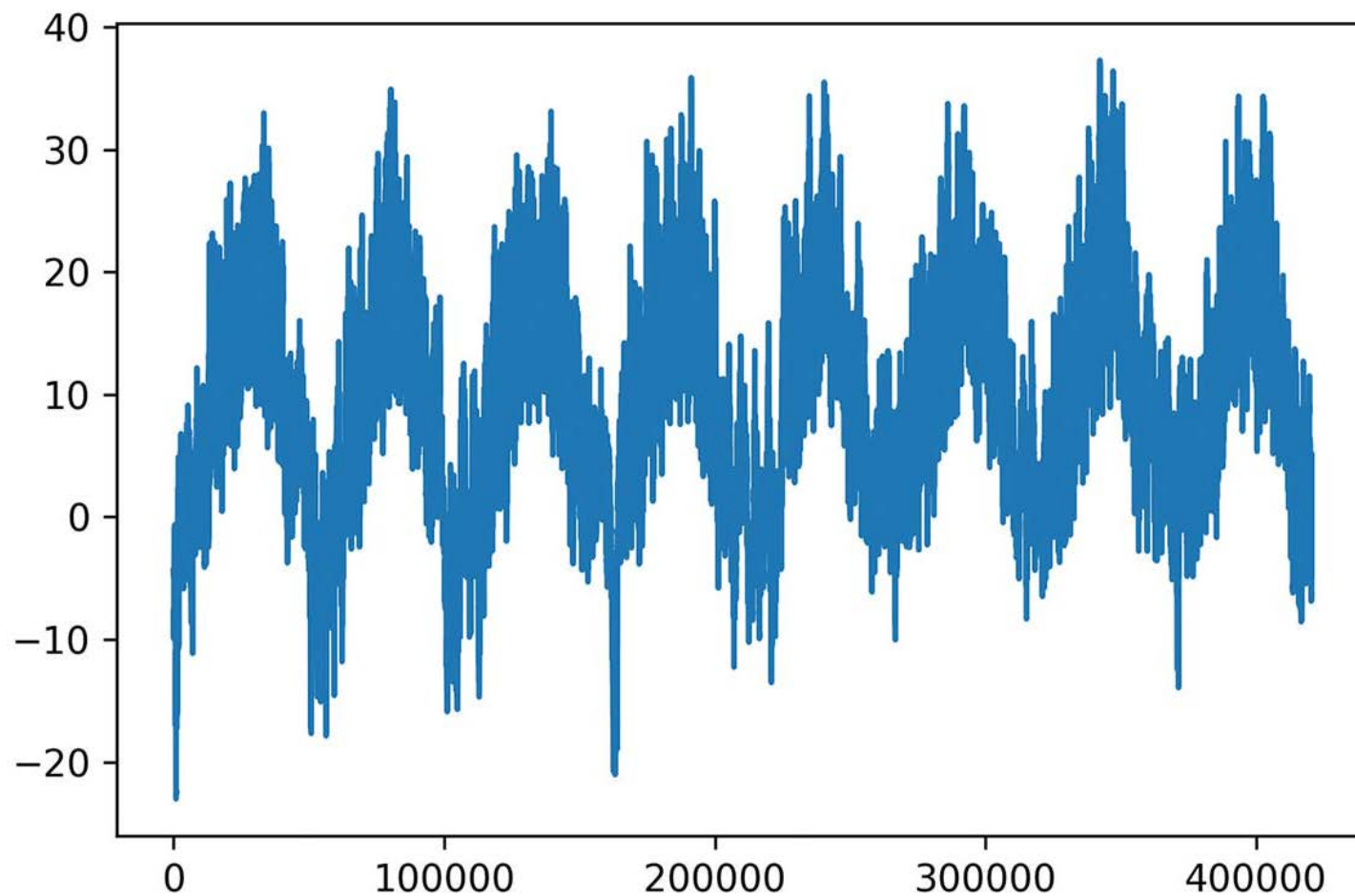
ВРЕМЕННЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

- **Временной последовательностью** могут быть любые данные, полученные с помощью измерений через равные промежутки времени.
- Временные последовательности присутствуют повсюду, независимо от того, смотрим ли мы на природные явления или модели человеческой деятельности.
- В отличие от других типов данных, работа с временными последовательностями включает в себя понимание **динамики** системы — ее периодических циклов, того, как она меняется с течением времени, ее регулярного режима и ее внезапных всплесков.

НАИБОЛЕЕ РАСПРОСТРАНЕННЫЕ ЗАДАЧИ, СВЯЗАННЫЕ С ВРЕМЕННЫМИ ПОСЛЕДОВАТЕЛЬНОСТЯМИ

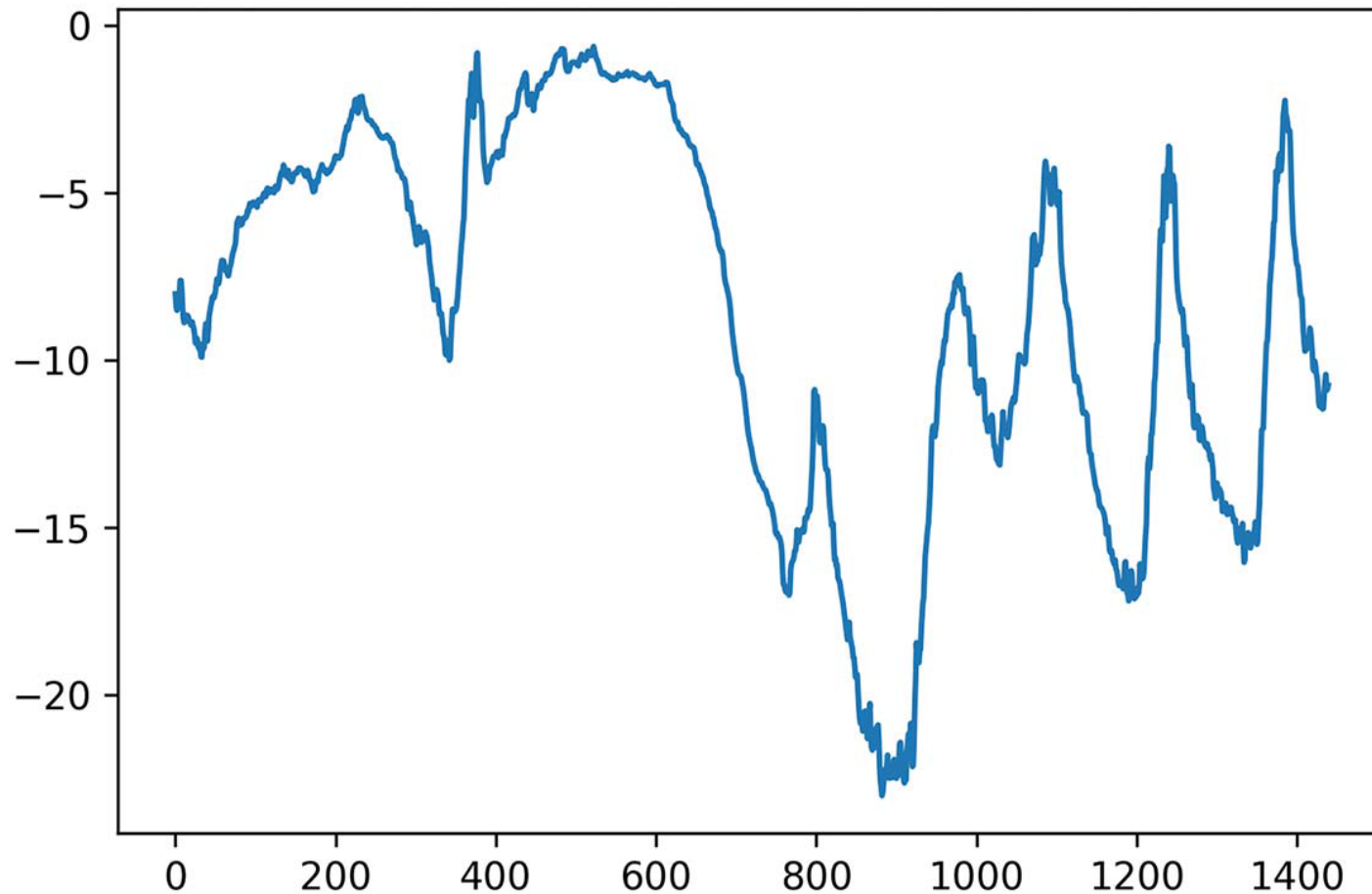
- **Прогнозирование** — прогнозирование того, что будет дальше в последовательности
- **Классификация** — присвоение одной или нескольких категориальных меток временным рядам
- **Обнаружение событий** — определение возникновения определенного ожидаемого события в непрерывном потоке данных
- **Обнаружение аномалий** — обнаружение чего-либо необычного, происходящего в непрерывном потоке данных

ПРИМЕР: РЯД ТЕМПЕРАТУРНЫХ ИЗМЕРЕНИЙ



Температура в течение
нескольких лет (°C)
с 10-минутным шагом

ПРИМЕР: РЯД ТЕМПЕРАТУРНЫХ ИЗМЕРЕНИЙ

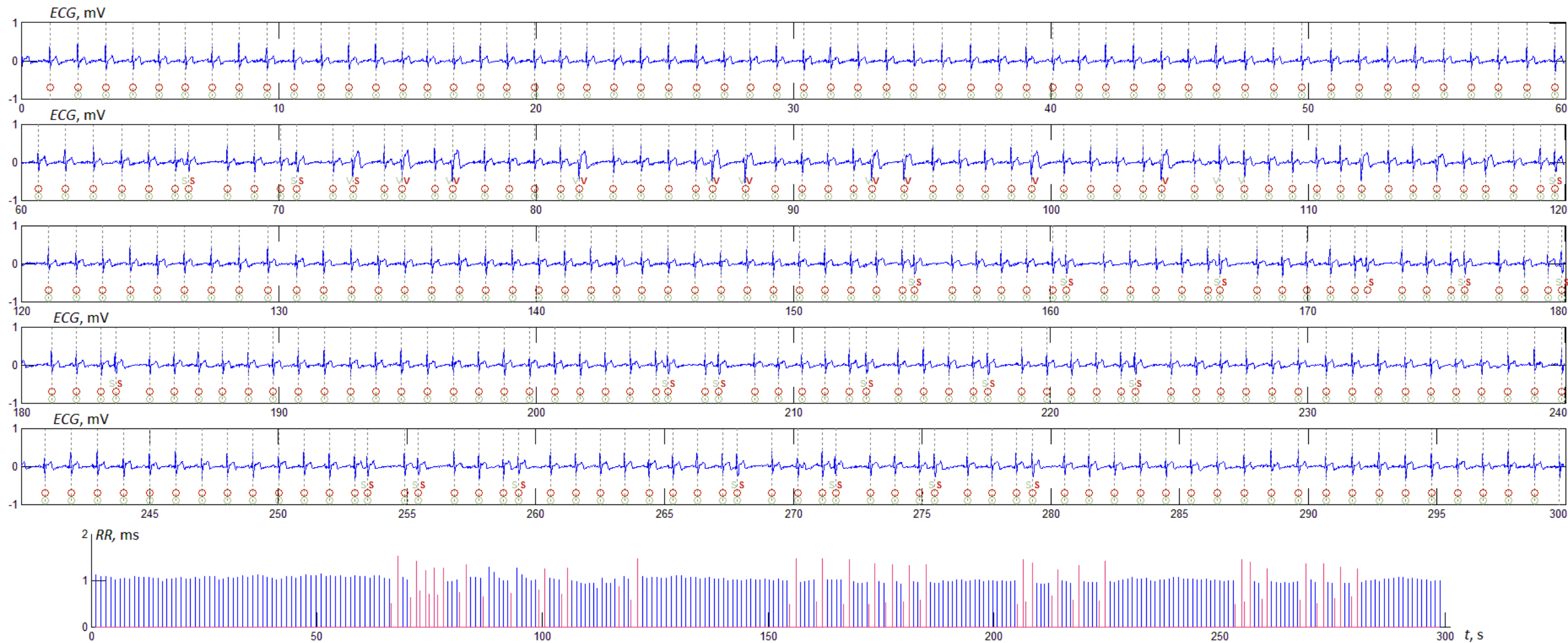


Температура в
течение первых 10
дней (°C)

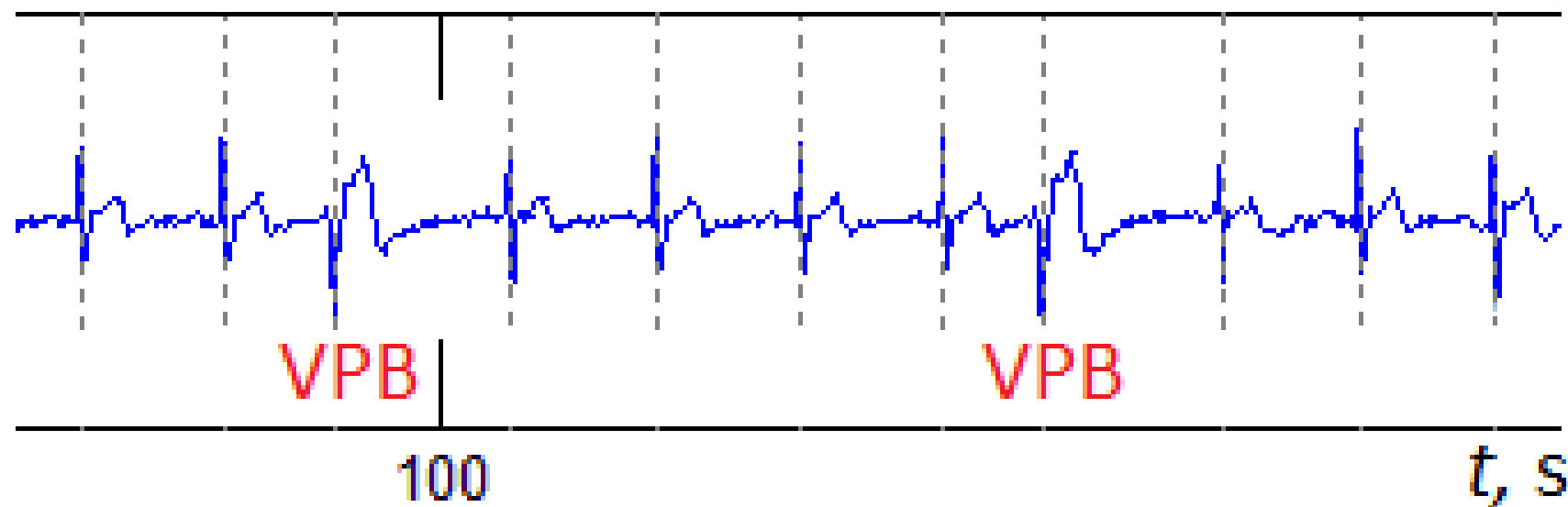
ПЕРИОДИЧНОСТЬ ВРЕМЕННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

- Всегда необходимо анализировать наличие **периодичности** в данных
- Периодичность в **нескольких временных масштабах** является важным и очень распространенным свойством временных последовательностей.
- При изучении данных эти шаблоны очень важны

ПРИМЕР: 5-МИНУТНАЯ ЗАПИСЬ ЭКГ



АРИТМИЧЕСКИЕ СОБЫТИЯ В ЭКГ

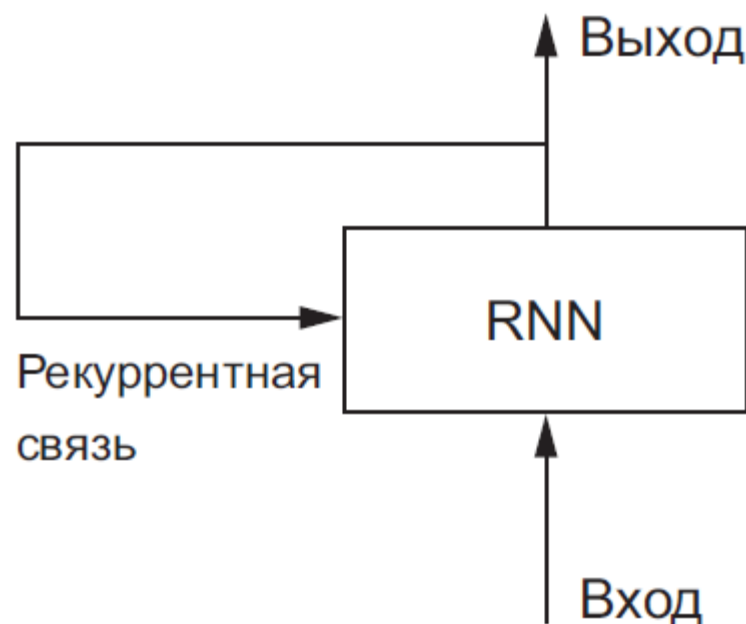


VPB – Ventricular Premature Beat (желудочковая экстрасистола)

РЕКУРРЕНТНАЯ НЕЙРОННАЯ СЕТЬ

- **Рекуррентная нейронная сеть** (recurrent neural network, RNN) обрабатывает последовательности, перебирая элементы последовательности и поддерживая состояние, содержащее информацию относительно того, что она «видела» до этого
- По сути, RNN — это тип нейронной сети, которая имеет внутреннюю петлю **обратной связи**
- Состояние RNN сбрасывается между обработкой двух разных независимых последовательностей, поэтому по-прежнему считаете одна последовательность по-прежнему считается единой точкой данных: одним входом в сеть. Эта точка данных больше не обрабатывается за один шаг; вместо этого сеть внутренне зацикливается на элементах последовательности

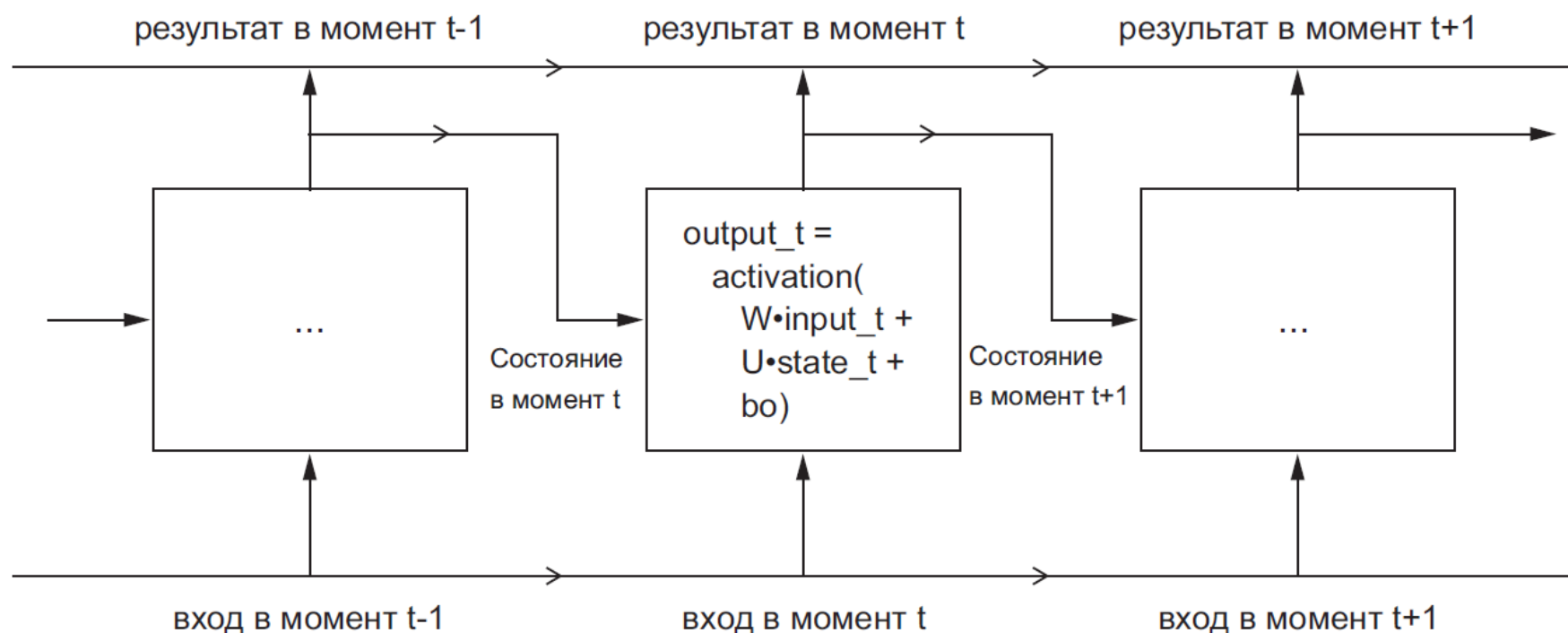
РЕКУРРЕНТНАЯ НЕЙРОННАЯ СЕТЬ



Рекуррентная
нейронная
сеть (RNN): сеть с петлей
обратной связи

ПРОСТАЯ СТРУКТУРА RNN

RNN – это цикл, который повторно использует величины, вычисленные во время предыдущей итерации цикла



Простая структура RNN, развернутая по времени

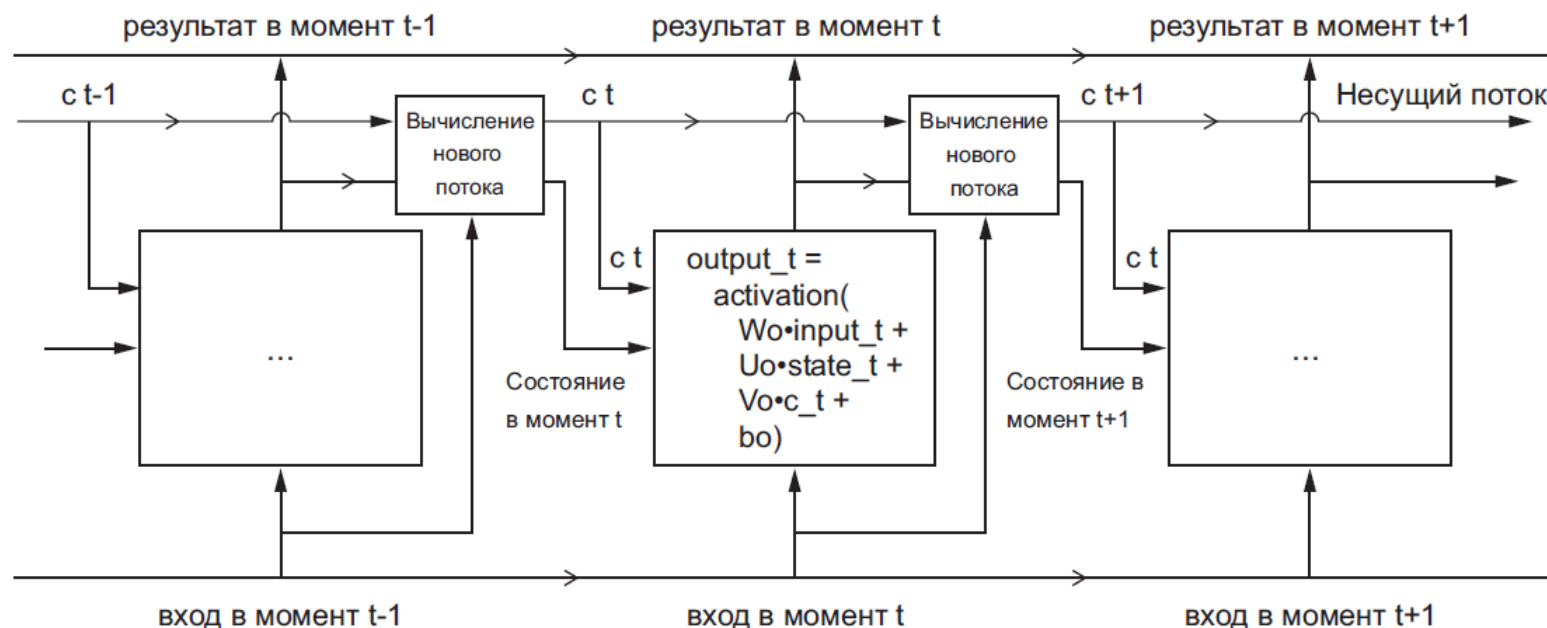
W_0 – весовая матрица для входных данных

U_0 – весовая матрица для данных состояния

b_0 – постоянная

СТРУКТУРА LSTM

Long Short-Term Memory (LSTM, долгая краткосрочная память) сохраняет информацию для последующего использования, тем самым предотвращая постепенное затухание старых сигналов во время обработки



Переход от простой RNN к LSTM: добавление несущего потока

W_o – весовая матрица для входных данных

U_o – весовая матрица для данных состояния

V_o – весовая матрица для данных несущего потока

b_o – постоянная

РАСШИРЕННЫЕ ВОЗМОЖНОСТИ RNN

- **рекуррентное прореживание** — особый встроенный способ использования прореживания для борьбы с переобучением в рекуррентных слоях
- **наложение рекуррентных слоев** — способ увеличения репрезентативности сети (за счет увеличения объема вычислений)
- **двунаправленные рекуррентные слои** — представляют одну и ту же информацию в рекуррентной сети разными способами, повышая точность и ослабляя проблемы, связанные с забыванием

ИСПОЛЬЗОВАНИЕ РЕКУРРЕНТНОГО ПРОРЕЖИВАНИЯ ДЛЯ БОРЬБЫ С ПЕРЕОБУЧЕНИЕМ

Правильный способ использования прореживания с RNN:

- На каждом шаге следует применять одну и ту же маску прореживания (один и тот же шаблон прореживаемых элементов) вместо маски прореживания, которая случайным образом изменяется от одного временного шага к другому
- Более того, для регуляризации представлений, образованных рекуррентными слоями, к внутренним периодическим активациям слоя должна применяться постоянная по времени маска прореживания
- Использование одной и той же маски прореживания на каждом временном шаге позволяет сети правильно распространять свою ошибку обучения во времени (случайная маска прореживания исказит сигнал ошибки и нарушит процесс обучения)

НАЛОЖЕНИЕ НЕСКОЛЬКИХ РЕКУРРЕНТНЫХ СЛОЕВ ДРУГ НА ДРУГА

Общая идея:

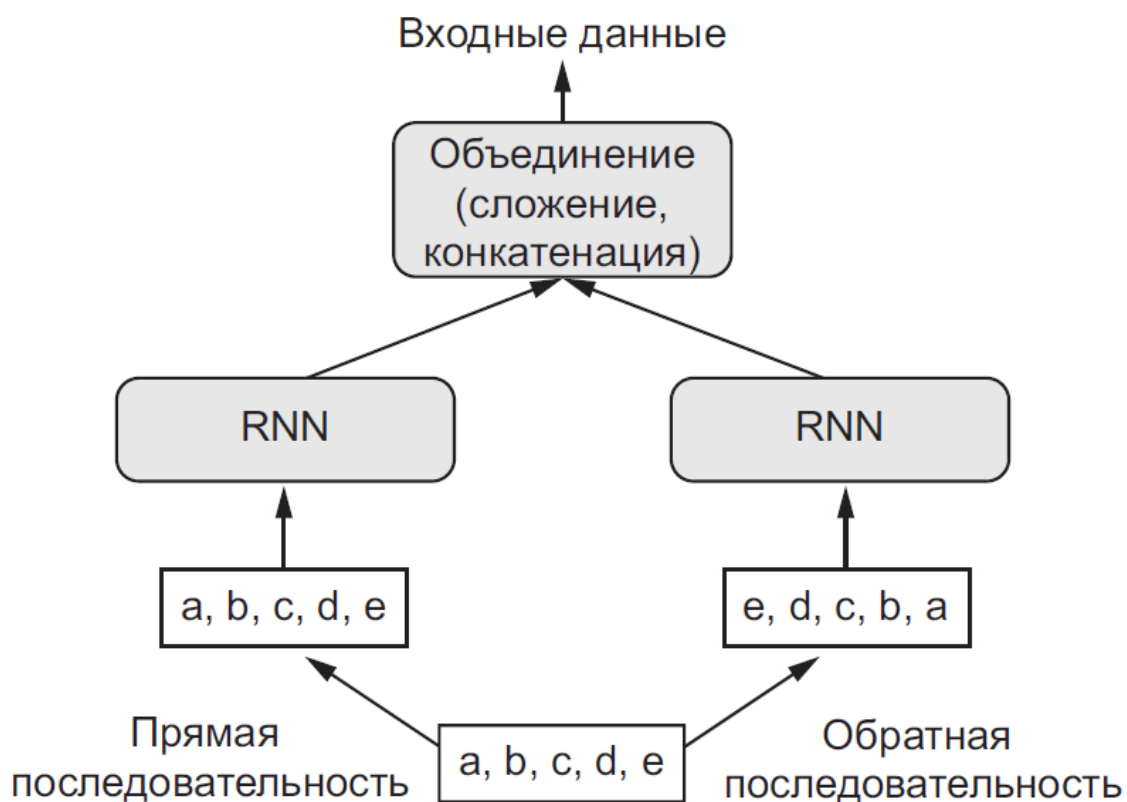
- Увеличивать емкость сети до тех пор, пока переобучение не станет основным препятствием

Увеличение емкости сети обычно осуществляется за счет увеличения количества элементов в слоях или добавления большего количества слоев.

Как это сделать для RNN:

- Рекуррентное наложение слоев — это классический способ создания более мощных рекуррентных сетей

ИСПОЛЬЗОВАНИЕ ДВУНАПРАВЛЕННЫХ RNN



Двунаправленная RNN использует чувствительность RNN к порядку поступления данных: она состоит из использования двух обычных RNN, каждая из которых обрабатывает входную последовательность в одном направлении (хронологически и анти-хронологически), а затем объединяет их представления

Обработывая последовательность в обоих направлениях, двунаправленная RNN может улавливать паттерны, которые могут быть пропущены однонаправленным RNN

КАК ПОВЫСИТЬ ПРОИЗВОДИТЕЛЬНОСТЬ RNN

- Отрегулируйте количество элементов в каждом рекуррентном слое структуры с наложением слоев, а также долю прореживаний
- Отрегулируйте скорость обучения, используемую оптимизатором RMSprop, или попробуйте другой оптимизатор
- Попробуйте использовать стек плотных (Dense) слоев поверх повторяющегося слоя вместо одного плотного слоя
- Улучшите входные данные модели: попробуйте использовать более длинные или более короткие последовательности или другую частоту дискретизации