

## 1. Питання № 7 : Модель довгої короткострокової пам'яті типу LSTM.

Модель довгої короткострокової пам'яті (LSTM) є спеціалізованою рекурентною нейронною мережею, яка використовується в глибокому навчанні для обробки послідовностей даних з великими інтервалами між взаємодіями. Вона була спроектована для подолання проблем короткострокової пам'яті та зниклого градієнта, які виникають при тренуванні традиційних рекурентних нейронних мереж (RNN).

Основні компоненти LSTM включають в себе "ворота" або механізми фільтрації інформації, які дозволяють моделі визначати, яку інформацію слід зберігати та передавати далі, а яку ігнорувати. Ці ворота включають:

1. **Ворота забуття (Forget Gate):** Визначає, яку інформацію з попереднього стану пам'яті слід забути або ігнорувати.
2. **Ворота введення (Input Gate):** Визначає, яку нову інформацію слід додати до пам'яті.
3. **Ворота виведення (Output Gate):** Визначає, який вихід слід виробляти на основі актуального стану пам'яті.

Ці ворота дозволяють LSTM зберігати та використовувати короткострокові та довгострокові залежності в послідовностях даних, забезпечуючи важливу структуру та здатність моделі робити передбачення на основі контексту.

У результаті використання LSTM модель може ефективно працювати з різноманітними завданнями, такими як машинне переклад, розпізнавання мови, генерація тексту та інші, де важлива обробка послідовностей даних з великими залежностями між елементами.

## 2. Питання № 9: Архітектура GRU, конструктивні параметри, особливості навчання.

Архітектура GRU (Gated Recurrent Unit):

GRU є іншим типом рекурентної нейронної мережі, яка подібна до LSTM, але має менше параметрів і, за теорією, може вирішити проблему зниклого градієнта більш ефективно. Вона складається з одного шару з одиниць пам'яті, а не з трьох, як у LSTM. Основною ідеєю GRU є використання двох воріт: ворота оновлення та ворота скидання.

1. **Ворота оновлення (Update Gate):** Відповідає за вирішення того, яка частина попереднього стану пам'яті має бути оновлена.
2. **Ворота скидання (Reset Gate):** Визначає, яка частина попереднього стану пам'яті має бути ігнорована.

Основна різниця між GRU і LSTM полягає в тому, що GRU має менше параметрів і воріт, але все одно може зберігати та передавати довгострокові залежності в послідовності даних.

### **Конструктивні параметри GRU:**

- Кількість одиниць пам'яті (units): Кількість нейронів у внутрішньому стані моделі.
- Функції активації: Зазвичай використовуються сигмоїда та тангенс гіперболічний.
- Ваги входів та ваги зворотного зв'язку: Параметри, які оптимізуються під час тренування для вирішення конкретного завдання.

### **Особливості навчання:**

- Зниклий градієнт: GRU був спроектований для вирішення проблеми зниклого градієнта, що може виникнути при тренуванні глибоких рекурентних мереж.
- Менше параметрів: У порівнянні з LSTM, GRU має менше параметрів, що може робити його привабливим в умовах обмежених ресурсів.
- Відмінна продуктивність для деяких завдань: GRU може показувати аналогічну або навіть кращу продуктивність, ніж LSTM, для деяких послідовних завдань, особливо в випадках обмежених обсягів даних для тренування.

Особливості навчання GRU включають в себе механізми автоматичного вагового оновлення, які дозволяють мережі адаптуватися до завдань з великою кількістю даних. Також, завдяки наявності гейтів, GRU може легше передавати та забувати інформацію на різних етапах часового ряду, що робить її більш придатною для моделювання довготривалих залежностей в порівнянні з традиційними RNN.

Через меншу кількість параметрів, ніж LSTM, GRU може бути менш витратною за ресурси для тренування та використання. Однак, вибір між LSTM і GRU часто залежить від конкретної задачі та доступних ресурсів.