



ОНЛАЙН-ОБРАЗОВАНИЕ

Автокодировщики

Жмем и преобразуем данные

Артур Кадулин
Преподаватель

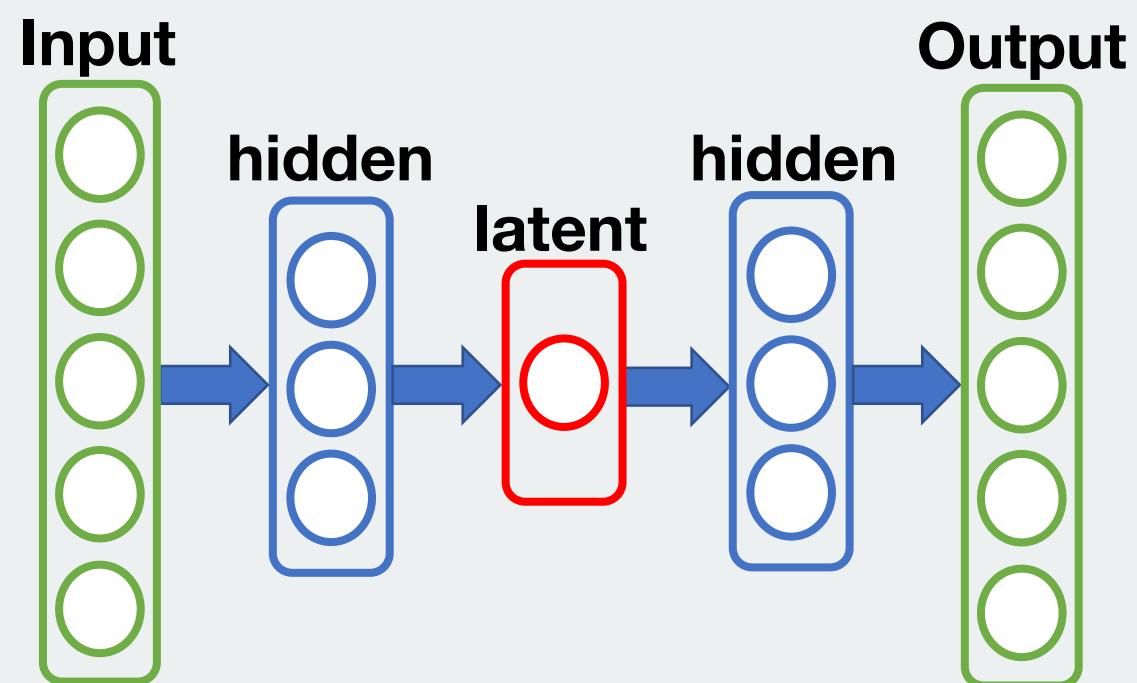


План на сегодня

1. **Что такое Автокодировщик**
2. Для чего нужны Автокодировщики и какими они бывают
3. Разреженный Автокодировщик
4. Практика: Автокодировщики
5. Практика: Бонус



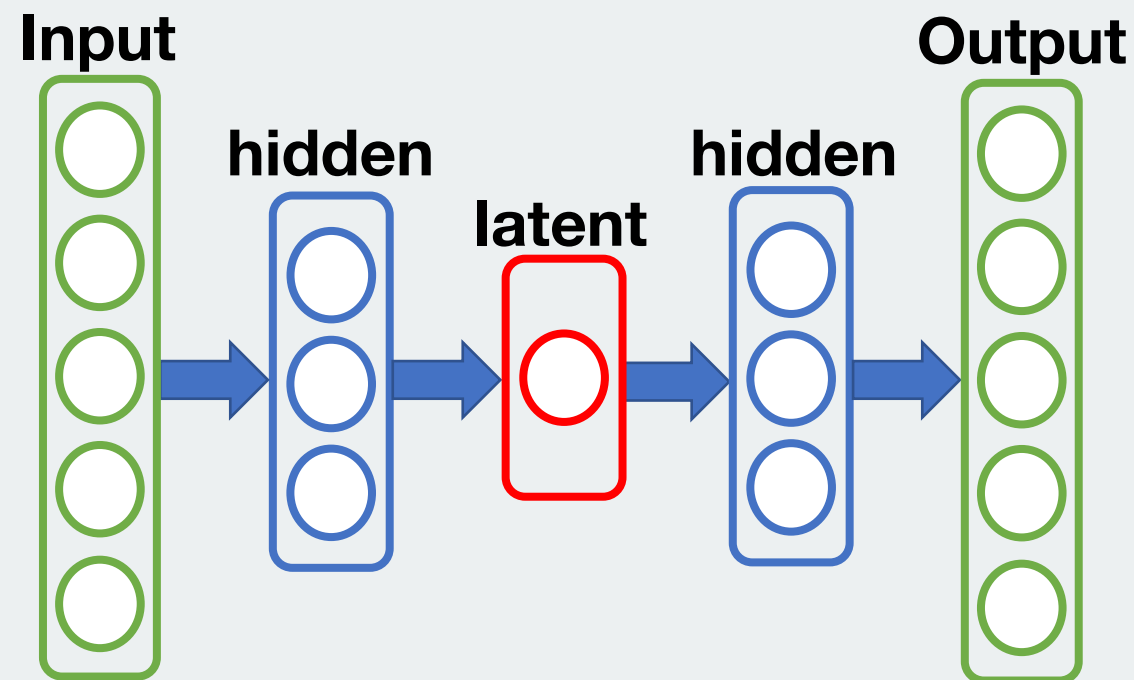
АВТОКОДИРОВЩИК



$$AE(x) = D(E(x))$$



АВТОКОДИРОВЩИК



$$AE(x) = D(E(x))$$

$$\mathcal{L}(x) = d(x, D(E(x)))$$



План на сегодня

1. Что такое Автокодировщик
- 2. Для чего нужны Автокодировщики и какими они бывают**
3. Разреженный Автокодировщик
4. Практика: Автокодировщики
5. Практика: Бонус



Автокодировщик

1. Сжатие данных:

Для описания одного объекта в пространстве грейскейл изображений размером 28x28 пикселей необходимо 784 байта. Однако 99.9...% объектов в этом пространстве не являются изображениями цифр. Говорят, что в этом 784-мерном пространстве есть многообразие рукописных цифр.



Автокодировщик

1. Сжатие данных
2. Предобучение сети

Если мы научились эффективно переводить объекты из пространства большей размерности в пространство меньшей размерности, значит мы научились извлекать признаки сильно связанные с самими объектами.



Автокодировщик

1. Сжатие данных
2. Предобучение сети
3. Переход в другое пространство

Если мы умеем отображать объекты из одного пространства в другое, почему бы не наложить дополнительные ограничения на второе пространство?



Автокодировщик

1. Сжатие данных
2. Предобучение сети
3. Переход в другое пространство
4. Восстановление данных

Если мы знаем как устроено пространство объектов и у нас есть только часть информации об одном из них, мы можем найти наиболее похожий



Автокодировщик

1. Типичные функции потерь для автокодировщиков

???



Автокодировщик

1. Типичные функции потерь для автокодировщиков

Любое дифференцируемое расстояние в исходном пространстве объектов может быть функцией потерь. Чаще всего используются

Кросс-энтропия

Среднеквадратичное отклонение

Косинусное расстояние



Автокодировщик

1. Типичные функции потерь для автокодировщиков
2. Ограничения на латентный слой

Ограничения на латентный слой можно наложить добавив слагаемое в функцию потерь.

Это может быть как дифференцируемая функция например L2 или KL

Так и ...?



Автокодировщик

1. Типичные функции потерь для автокодировщиков
2. Ограничения на латентный слой

Ограничения на латентный слой можно наложить добавив слагаемое в функцию потерь.

Это может быть как дифференцируемая функция например L2 или KL

Так и еще одна нейросеть!



Автокодировщик

1. Типичные функции потерь для автокодировщиков
2. Ограничения на латентный слой

Ограничения на латентный слой можно наложить добавив слагаемое в функцию потерь.

Это может быть как дифференцируемая функция например L2 или KL

Так и еще одна нейросеть!

Вообще, идеей использования отдельной нейросети в качестве компонента функции ошибки мы будем пользоваться очень много



Автокодировщик

1. Типичные функции потерь для автокодировщиков
2. Ограничения на латентный слой
3. Аугментация данных

Что, если на вход мы будем подавать зашумленные картинки, а на выходе будем требовать восстановить исходные?



Автокодировщик

1. Типичные функции потерь для автокодировщиков
2. Ограничения на латентный слой
3. Аугментация данных

Что, если на вход мы будем подавать зашумленные картинки, а на выходе будем требовать восстановить исходные?

Оказывается это не только возможно, но и помогает при обучении!



План на сегодня

1. Что такое Автокодировщик
2. Для чего нужны Автокодировщики и какими они бывают
- 3. Разрезанный Автокодировщик**
4. Практика: Автокодировщики
5. Практика: Бонус



Разреженный Автокодировщик

Давайте, для примера, наложим на латентный слой автокодировщика ограничение, минимизирующее количество «активных» нейронов в нем.



Разреженный Автокодировщик

Давайте, для примера, наложим на латентный слой автокодировщика ограничение, минимизирующее количество «активных» нейронов в нем.

Допустим каждый нейрон либо активен, либо — нет. Пусть тогда вероятность быть активным равна ρ



Разреженный Автокодировщик

Давайте, для примера, наложим на латентный слой автокодировщика ограничение, минимизирующее количество «активных» нейронов в нем.

Допустим каждый нейрон либо активен, либо — нет. Пусть тогда вероятность быть активным равна ρ

$$\mathcal{L}(x) = d\left(x, D(E(x))\right) + D_{KL}(\rho \| E(x))$$



Разреженный Автокодировщик

$$E(x) = \hat{\rho}$$

$$D_{KL}(\rho || E(x)) = \sum_i \rho_i \log \frac{\rho_i}{\hat{\rho}_i} = \dots ?$$



Разреженный Автокодировщик

$$E(x) = \hat{\rho}$$

$$D_{KL}(\rho \| E(x)) = \sum_i \rho_i \log \frac{\rho_i}{\hat{\rho}_i} = \rho \log \frac{\rho}{\hat{\rho}} + (1 - \rho) \log \frac{(1 - \rho)}{(1 - \hat{\rho})}$$



Разреженный Автокодировщик

$$E(x) = \hat{\rho}$$

$$D_{KL}(\rho \| E(x)) = \sum_i \rho_i \log \frac{\rho_i}{\hat{\rho}_i} = \rho \log \frac{\rho}{\hat{\rho}} + (1 - \rho) \log \frac{(1 - \rho)}{(1 - \hat{\rho})}$$

$$\rho \log \frac{\rho}{\hat{\rho}} = \rho \log \rho - \rho \log \hat{\rho}$$



Разреженный Автокодировщик

$$E(x) = \hat{\rho}$$

$$D_{KL}(\rho \| E(x)) = \sum_i \rho_i \log \frac{\rho_i}{\hat{\rho}_i} = \rho \log \frac{\rho}{\hat{\rho}} + (1 - \rho) \log \frac{(1 - \rho)}{(1 - \hat{\rho})}$$

$$\rho \log \frac{\rho}{\hat{\rho}} = \rho \log \rho - \rho \log \hat{\rho}$$

$$\min_E D_{KL}(\rho \| E(x)) = \min_E (-\rho \log \hat{\rho} - (1 - \rho) \log(1 - \hat{\rho}))$$



План на сегодня

1. Что такое Автокодировщик
2. Для чего нужны Автокодировщики и какими они бывают
3. Разреженный Автокодировщик
4. **Практика: Автокодировщик**
5. **Практика: Бонус**





Спасибо
за внимание!