



УНИВЕРСИТЕТ ИТМО

Генерация изображений чисел

Санкт-Петербург
2022



~~Перечень клоунов~~ Состав команды

Щепчун Андрей

Шохов Максим

Малюкова Мария



План действий

Deadline	Project status
26/09/2022	Выбор темы и поиск литературы
05/10/2022	Изучение найденной литературы и поиск новой
15/10/2022	Представление темы
22/10/2022	Поиск датасетов
05/11/2022	Выбор подхода к обучению нейронных сетей и написание “чернового” кода



План действий

Deadline	Project status
15/11/2022	Презентация “чернового” варианта проекта
22/11/2022	Исправление кода в соответствии с комментариями преподавателя
11/12/2022	Исправление презентации в соответствии с комментариями преподавателя
15/12/2022	Итоговая презентация проекта



Распределение ролей

Написанием кода, по итогу, занимался Шохов.

Редактированием кода программы Щепчун.

Написанием комментариев и презентацией Малюкова.

Проблемы и методы решения

Проблема	Методы решения
Проблемы с набором данных	Проверка входных данных
	Проверка на случайных данных
	Проверка загрузчика данных
	Проверка взаимоотношения входа и выхода
	Проверка набора данных на шумы
	Изменение порядка набора данных
	Снижение несбалансирования классов
	Увеличение датасета
	Проверка партий



Проблемы и методы решения

Проблема	Методы решения
Проблемы аугментации	Откалибровка признаков
	Проверка аугментации данных
	Проверка предобработки
Проблемы реализации	Проверка функции потерь
	Проверка слоев
	Увеличение размера сети
	Поиск скрытых ошибок измерений

Проблемы и методы решения

Проблема	Методы решения
Проблемы обучения	Проверка на маленьком наборе данных
	Проверка инициализации весов
	Изменение гиперпараметров
	Уменьшение регуляризации
	Дать время на обучение
Проблема прокрастинации	Использование техник борьбы с прокрастинацией

Датасет

В результате командного штурма различных датасетов был выбран наиболее подходящий - mnist.

Метрики

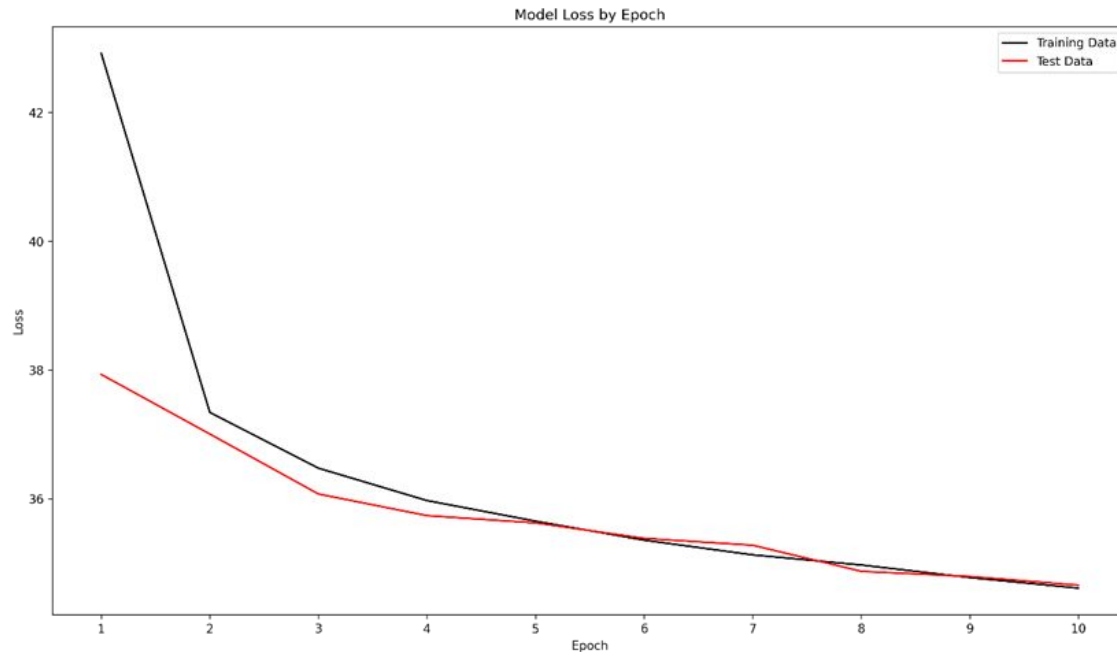
Из огромного множества метрик, наш выбор пал на mse.

Модель

В качестве модели были выбраны вариационные автокодеры (VAE)

Обучение

После различных попыток было решено, что 10 эпох обучения дают ощутимо меньшую ошибку



Результат работы кода

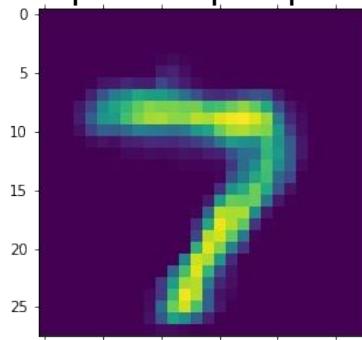
MNIST digit representation in the 2D Latent Space



Распределение чисел из набора данных в скрытом пространстве

Результат работы кода

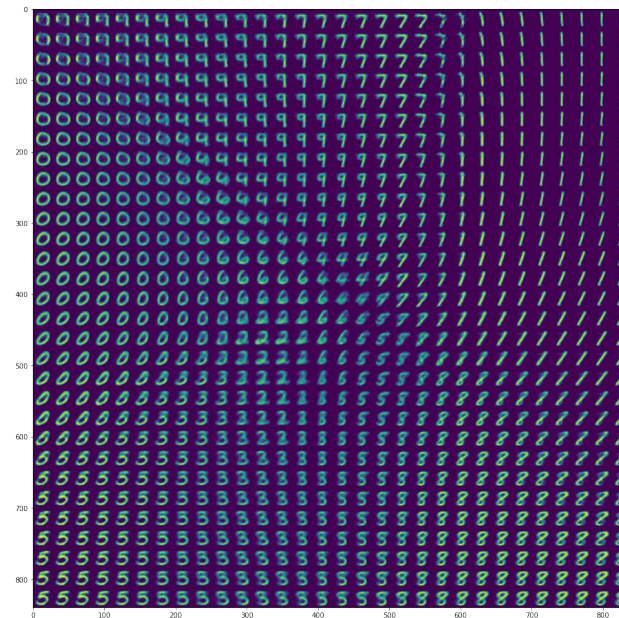
Подбирая значения мат ожидания и дисперсии – мы можем сгенерировать изображение желаемого числа. Например, параметры 0.5 и 1.5 будут соответствовать числу 7 на скрытом пространстве.



Число, соответствующее параметрам 0.5 и 1.5

Больше результатов

Так, на рисунке мы можем увидеть сгенерированные 30×30 чисел на скрытом пространстве с изменением параметра по оси X от 1.5 до -1.5 и по оси Y от -1.5 до 1.5.



Немного тестов

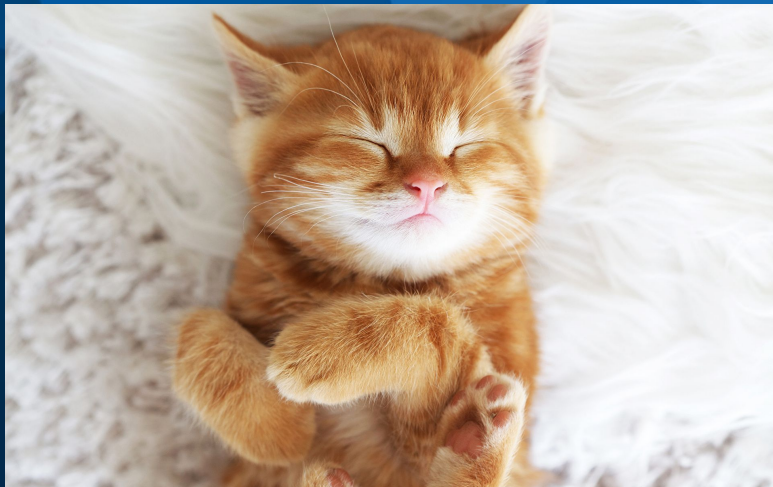
Для нейросети, распознающей рукописные символы были “проброшены” 900 сгенерированных чисел. Потери не были большими, однако и точность оказалась небольшой.

```
29/29 [=====] - 0s 622us/step - loss: 1.7320 - accuracy: 0.2733
```


Списки литературы

1. <https://habr.com/ru/company/sberbank/blog/586926/>
2. <https://intuit.ru/studies/courses/940/409/lecture/17868>
3. <https://portal.tpu.ru/SHARED/j/JBOLOTOVA/academic/ComputerGraphics/CGStudyBook.pdf>
4. “Zero-Shot Text-to-Image Generation”. Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, Ilya Sutskever [26/02/2021]
5. <https://habr.com/ru/post/334944/>
6. <http://blog.data.lytica.ru/2018/05/blog-post.html>

Спасибо за внимание!



IT's *MO*re than a
UNIVERSITY

котик наелся и жив, а мы - нет(