

Тема: Нейробайєсівський підхід у машинному навчанні на прикладі варіаційного автокодувальника

Виконав: студент групи КА-41
Меденцій В.А.

Науковий керівник : доцент кафедри ММСА ІПСА
канд.фіз-мат.наук Каніовська І.Ю.

Об'єкт, предмет і мета

Об'єкт: генеративна модель машинного навчання, що використовує байєсівський підхід

Предмет: методи відновлення розподілу даних, що дозволяють генерувати нові дані з нього

Мета: дослідити особливості байєсівського підходу у машинному навчанні, побудувати модель породження даних, проаналізувати результати роботи алгоритму при різних гіперпараметрах моделі

Актуальність

Генеративні моделі дозволяють вирішувати такі задачі

— створення нових даних для збільшення навчальної вибірки для побудови інших моделей машинного навчання

— відновлення роздільної здатності зображень та розфарбування

— тренування генеративної моделі дозволяє знайти представлення даних у просторі латентних змінних меншої розмірності

— виявлення аномальних даних та викидів

Постановка задачі

Побудувати модель для знаходження розподілу даних та породження нових даних

Провести порівняльний аналіз результатів в залежності від розмірності простору латентних змінних

Дослідити можливі модифікації методу та реалізувати умовну генеративну модель

Особливості байєсівського підходу

- використання апріорних знань дослідників
- гарно працюють на вибірках малого обсягу
- дозволяють вводити латентні змінні, що суттєво знижує кількість параметрів моделі
- вимагають довготривалі обчислювальні ресурси

Задача породження зображень

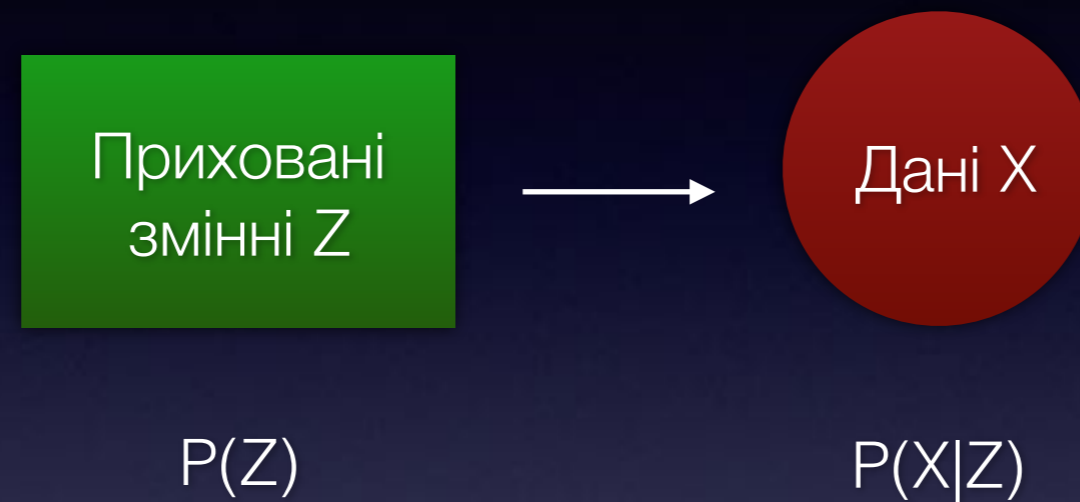
Дано: X - вибірка з 60000 чорно-білих зображень,
розмірності 28×28



Знайти розподіл $P(X|\theta)$

Модель варіаційного автокодувальника

Припущення 1: Дані породженні двоступінчатим процесом



Припущення 2: Максимізація нижньої варіаційної оцінки правдоподібності

$$\log P_{\theta}(X) \geq E_z \log P_{\theta}(X|Z) - KL(Q_{\phi}(Z|X) || P(Z))$$

$$\theta, \phi = \operatorname{argmax}_{\theta, \phi} \sum_i E_z (\log p_{\theta}(x_i|z)) - KL(q_{\phi}(z|x_i) || p_{\theta}(z))$$

Модель варіаційного автокодувальника

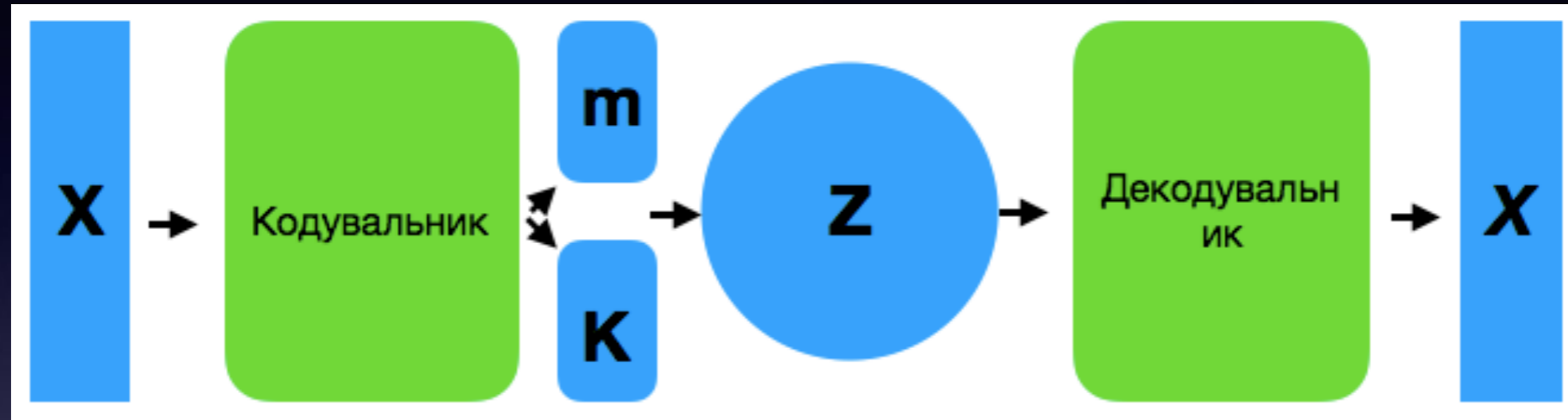
Припущення 3: $P(Z) \sim N(0; I)$

Припущення 4: $Q(Z|X, \Phi) \sim N(m; K)$, де K - діагональна матриця

Припущення 5: $P(X|Z, \theta)$ має розподіл Бернуллі

Модель варіаційного автокодувальника

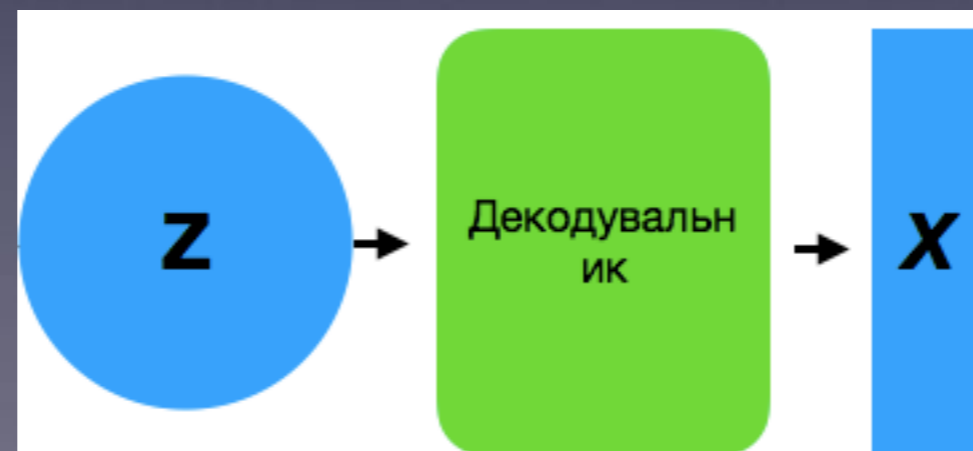
1. Знаходження розподілу



$Q(Z|X, \Phi) \sim$ Кодувальник

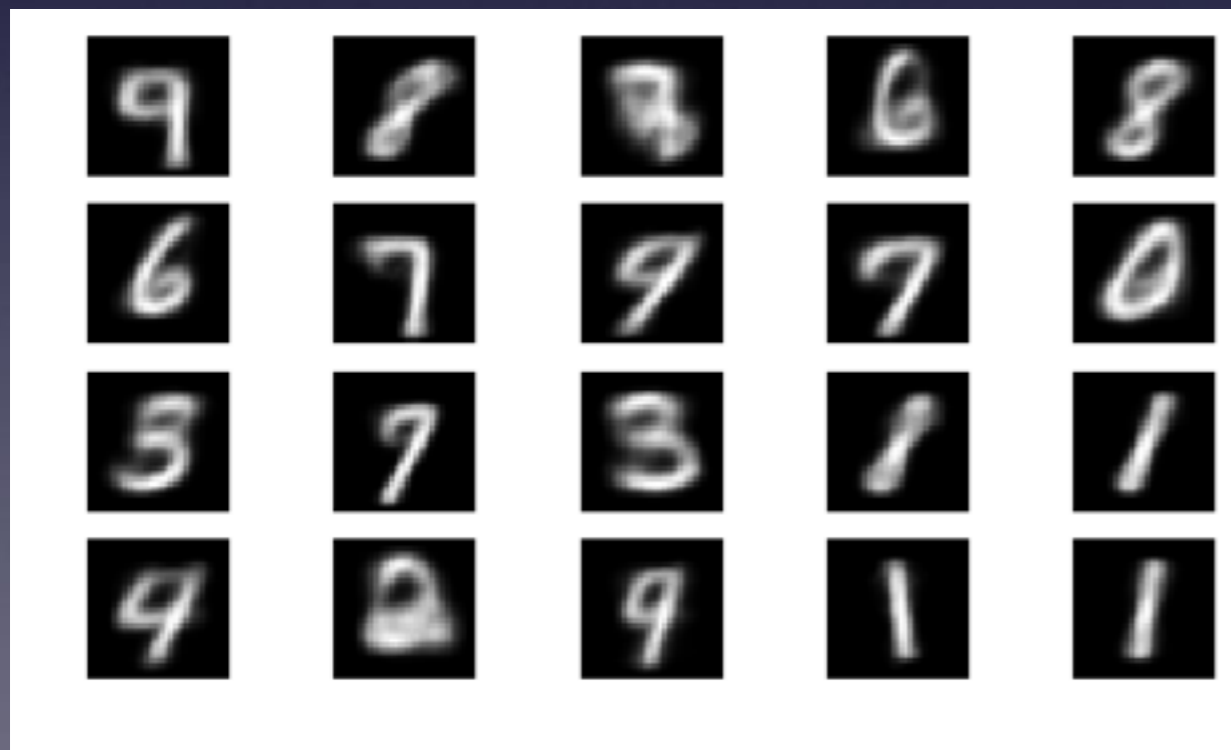
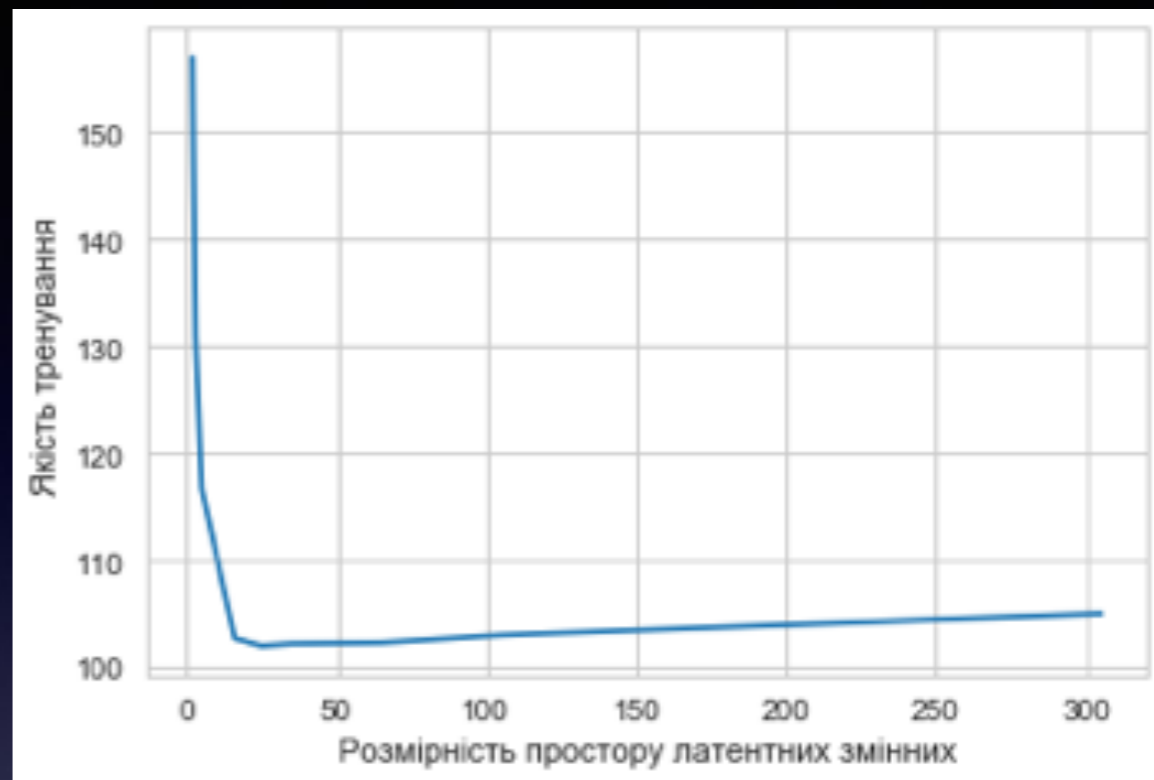
$P(X|Z, \theta) \sim$ Декодувальник

2. Генерація нових зображень

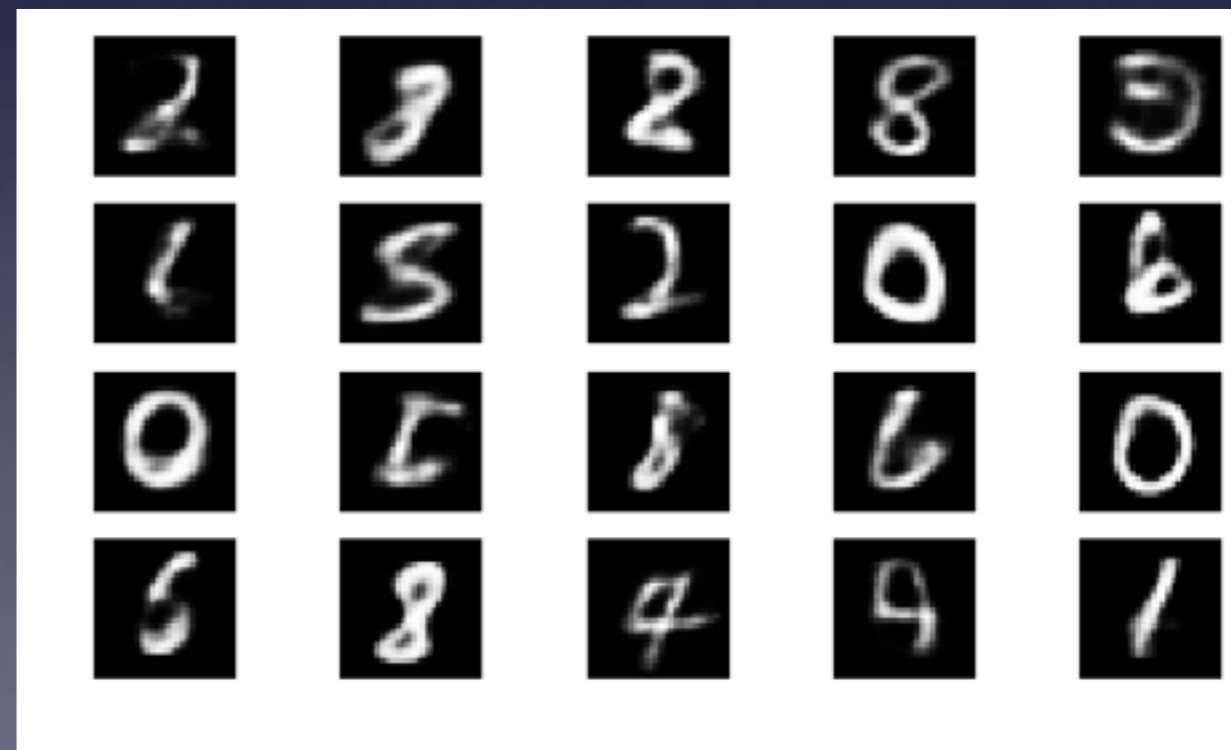


Генерація чорно-білих зображень

Якість навчання в залежності від розмірності простору латентних змінних



Розмірність = 25, якість = 102



Розмірність = 2, якість = 157

Навчання моделі

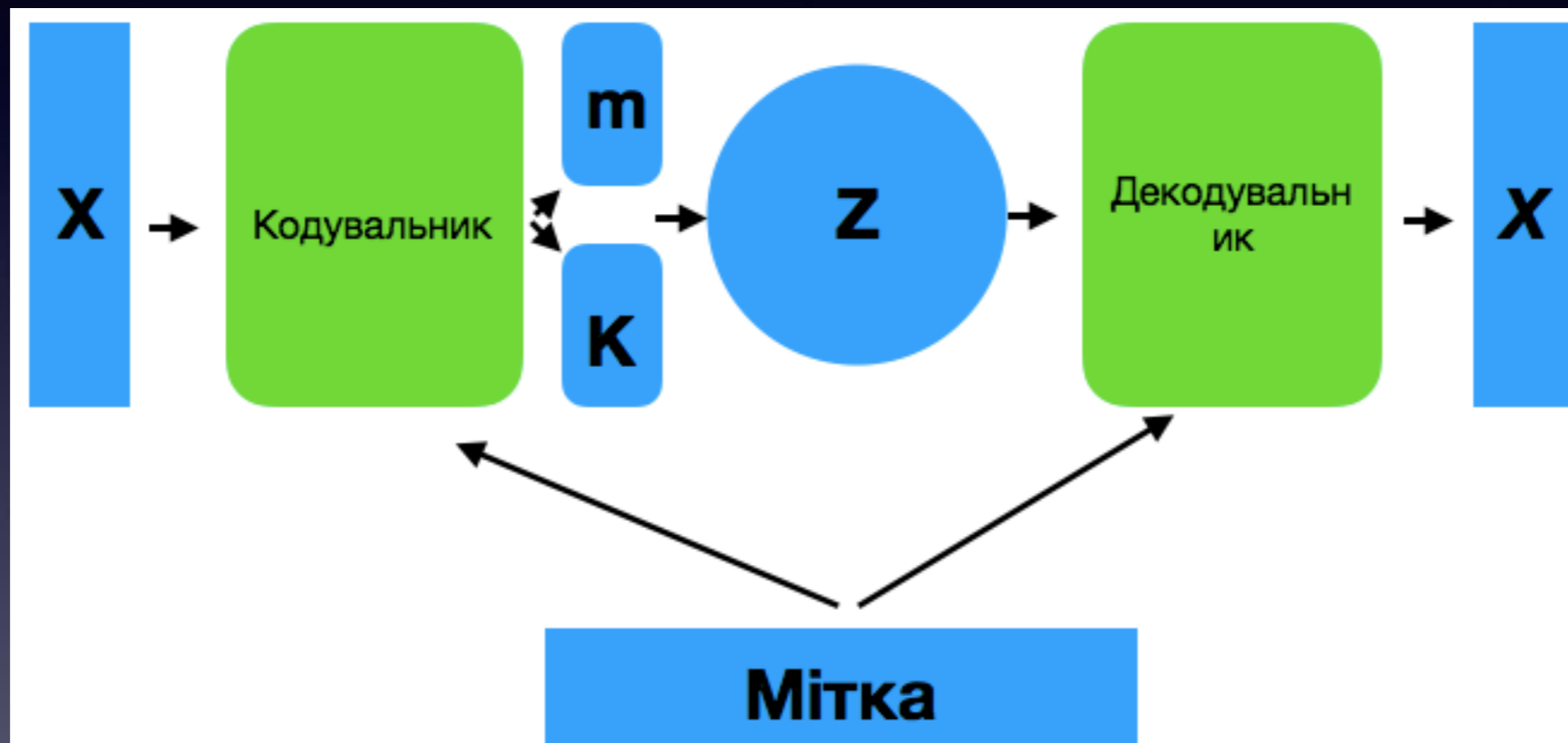
Кодувальник та декодувальник: трьох шарові нейронні мережі з ReLU активаційними функціями

Оптимізаційний алгоритм: RMSProp - модифікація стохастичного градієнтного спуску з адаптивним кроком навчання

Ініціалізація параметрів: метод Ксав'є - для уникнення «згасання» градієнтів

Умовний варіаційний автокодувальник

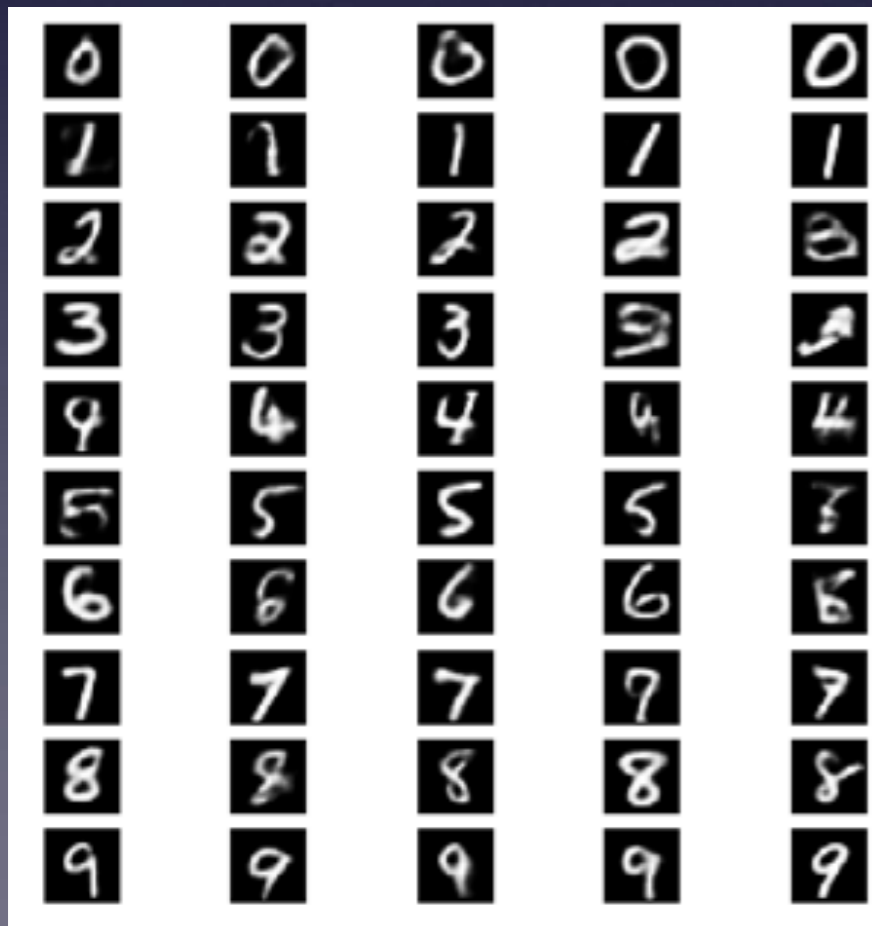
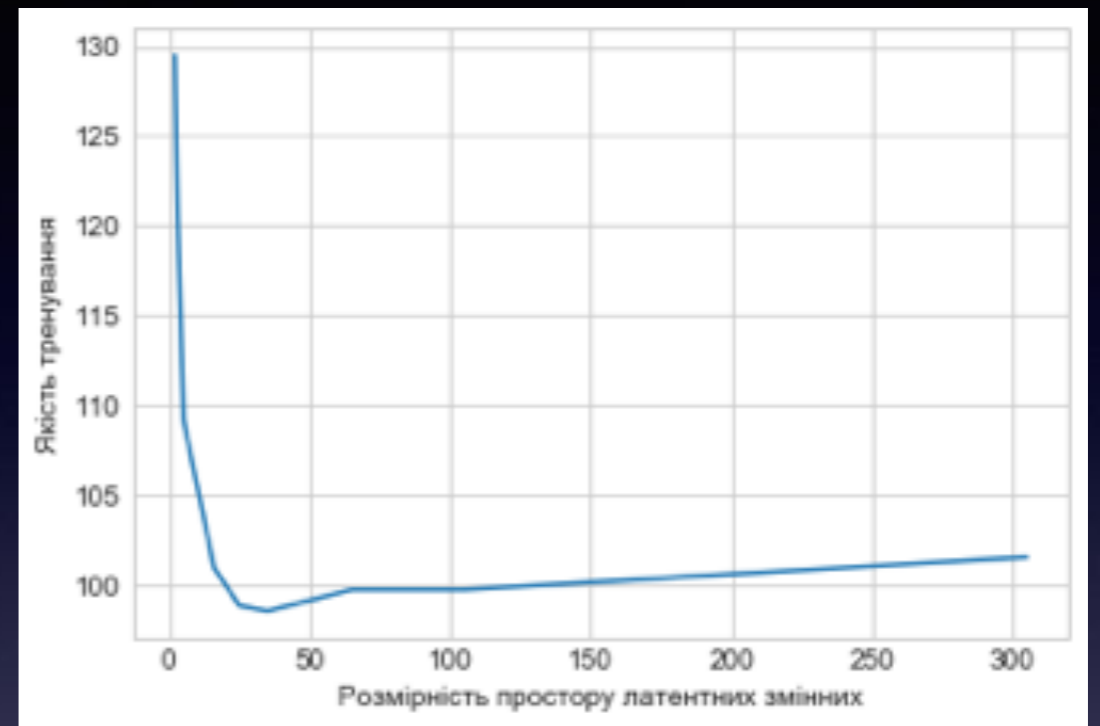
- дозволяє контролювати дані, що генеруються



- мітка визначає номер цифри, що генерується

Умовний варіаційний автокодувальник

Якість навчання в залежності від розмірності простору латентних змінних



Розмірність = 35, якість = 98.5

Модель варіаційного автокодувальника

Переваги

1. використання латентних змінних для представлення даних
2. наявність функціоналу якості моделі

Недоліки

1. зображення є менш чіткими порівняно з іншими алгоритмами

Можливі модифікації моделі

1. Варіаційний автокодуювальник, що дозволяє контролювати складність простору латентних змінних
2. Використання інших розподілів для моделювання простору латентних змінних, зокрема дискретних
3. Варіаційні автокодуювальники стійкі до шуму

Висновки

1. Реалізовані класична та умовна моделі варіаційного автокодуювальника можуть бути використані для породження нових даних при оптимальному значенні розмірності простору латентних змінних
2. Умовна модель має кращу якість, оскільки дає моделі додаткову інформацію
3. Подальші дослідження можуть бути проведені у напрямку модифікації моделі, які б враховували природу даних

Дякую за увагу