Generative Machine Learning
opportunities and challenges

Ekapol Chuangsuwanich
ekapolc@cp.eng.chula.ac.th

Department of Computer Engineering, Chulalongkorn University

26 May 2022
AI & IoTs Summit 2022

Slides: https://bit.ly/aiiotgenerative

About me

Lecturer at Chulalongkorn University

Research focus: ASR, NLP, Bioinformatics, etc.

Various industry collaborations

Generative Machine Learning?

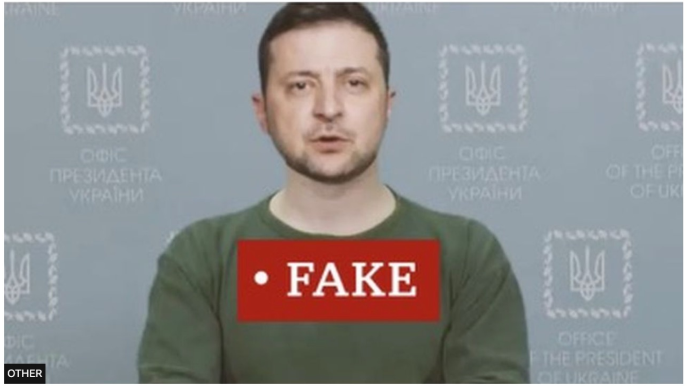
- Models that can learn the distribution of the data
  - Can be used to generate
    - Pictures, text, and more!
  - Multiple different algorithms
    - Explicit or implicit learning
      - Mostly deep-learning-based

Deepfake presidents used in Russia-Ukraine war

By Jane Wakefield
BBC Technology

18 March

Russia-Ukraine war



The deepfake appeared on the hacked website of Ukrainian TV network Ukrayina 24

Source: https://www.bbc.com/news/technology-60780142

Agenda

- Examples of generative modeling
    - Housing product insights
    - X-ray image description
    - Text-to-speech

ML for product development

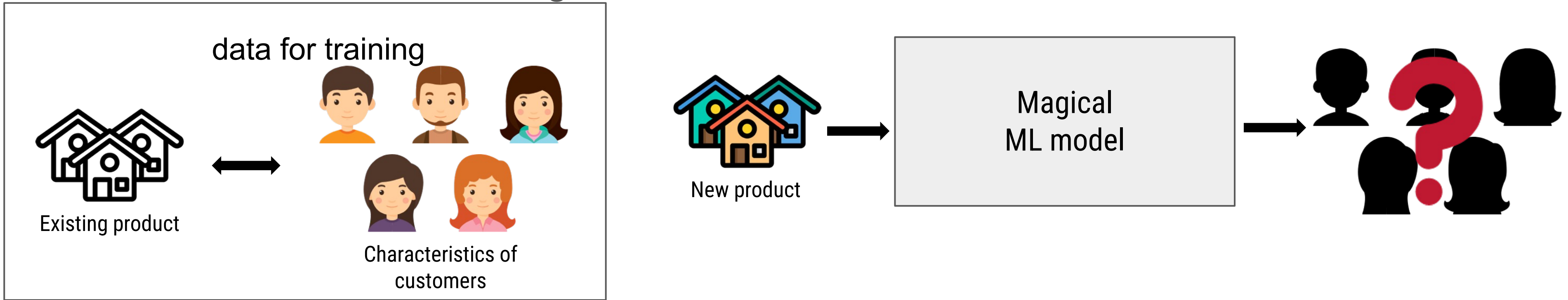
Work in colloboration with Home buyer's Group
http://home.co.th

For real estates, no two products are the same.

Development is based on gut feeling.

Can we make some educated guess about a new product

- popularity
- the type of potential buyers
- whether to add or remove some features
- the best marketing channel



ML for product development

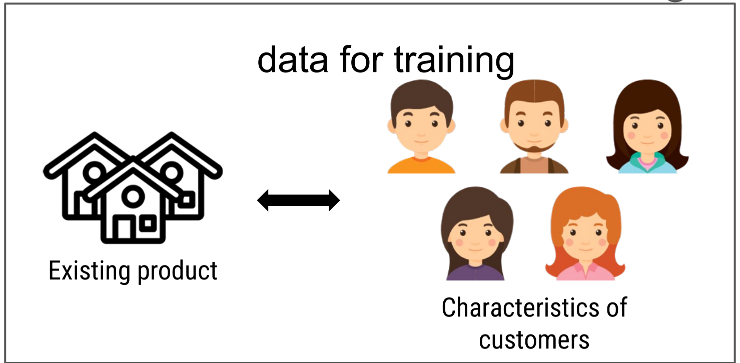
For real estates, no two products are the same.

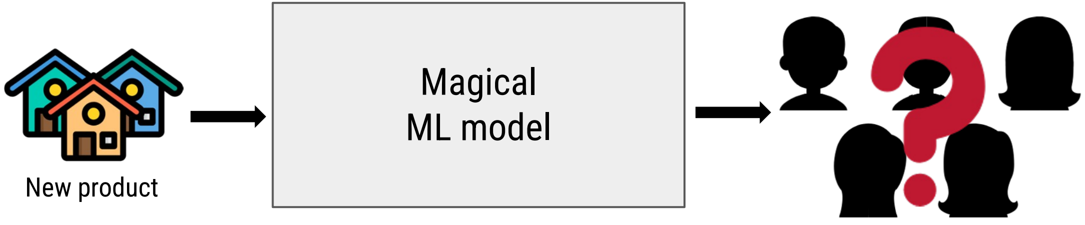
Development is based on gut feeling.

Can we make some educated guess about a new product

- popularity
- the type of potential buyers
- whether to add or remove some features
- the best marketing channel

We want to learn the distribution of the user given some input.
How?



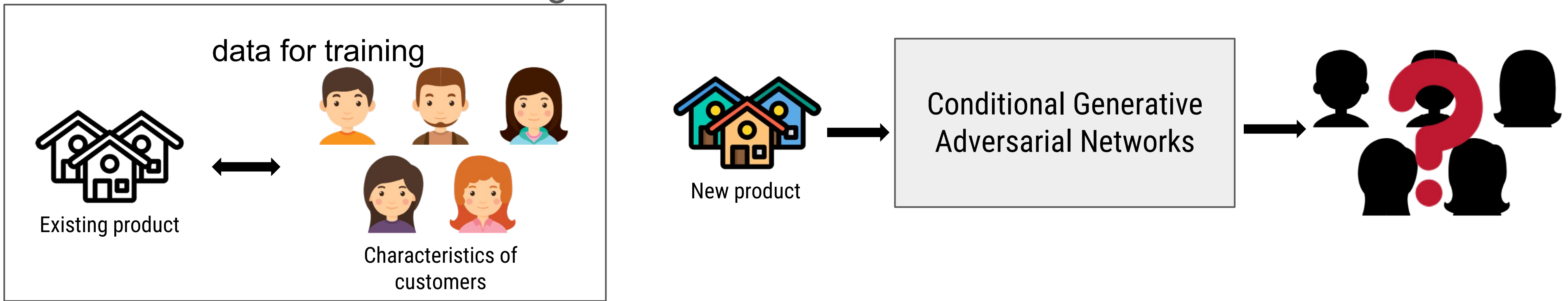


ML for product development

For real estates, no two products are the same.
Development is based on gut feeling.
Can we make some educated guess about a new product

- popularity
- the type of potential buyers
- whether to add or remove some features
- the best marketing channel

GAN!



Generative Adversarial Networks (GANs)

0.1, -0.3, .. → Generator →  → Discriminator → Real or Fake

z

x

Y=f(x)

Consider a money counterfeiter

- He wants to make fake money that looks real
- There's a police that tries to differentiate fake and real money.

The counterfeiter is the adversary and is generating fake inputs. – Generator network

The police is try to discriminate between fake and real inputs. – Discriminator network

Generative Adversarial Networks (GANs)

0.1, -0.3, .. → Generator →  → Discriminator → Real

z

x

Generative Adversarial Networks (GANs)

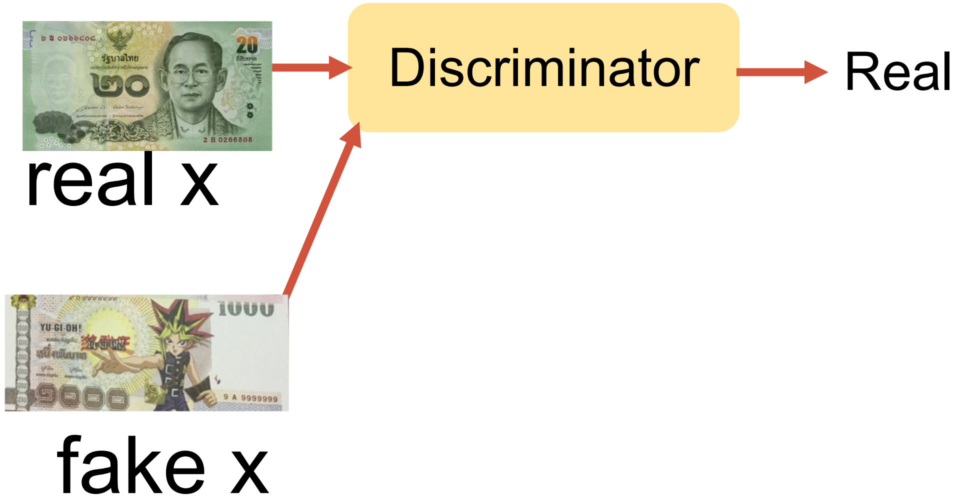
0.1, -0.3, .. → Generator → 

→ Discriminator → Fake

z

x

Generative Adversarial Networks (GANs)



Generative Adversarial Networks (GANs)

0.1, -0.3, .. → Generator →  → Discriminator → Fake

z

x

gradient

The generator learns to be better by
the gradient given by the discriminator

Conditional GAN (CGAN)

0.1, -0.3, ..
z

1.0
c

Generator

x

1.0
c

Discriminator

Real or Fake

Y=f(x,c)

GAN can be conditioned (controlled) to generate things you want by concatenating additional information

Conditional GAN (CGAN)

0.1, -0.3, .. → Generator → 

→ Discriminator → Real or Fake

z

-1.0

c

x

-1.0

c

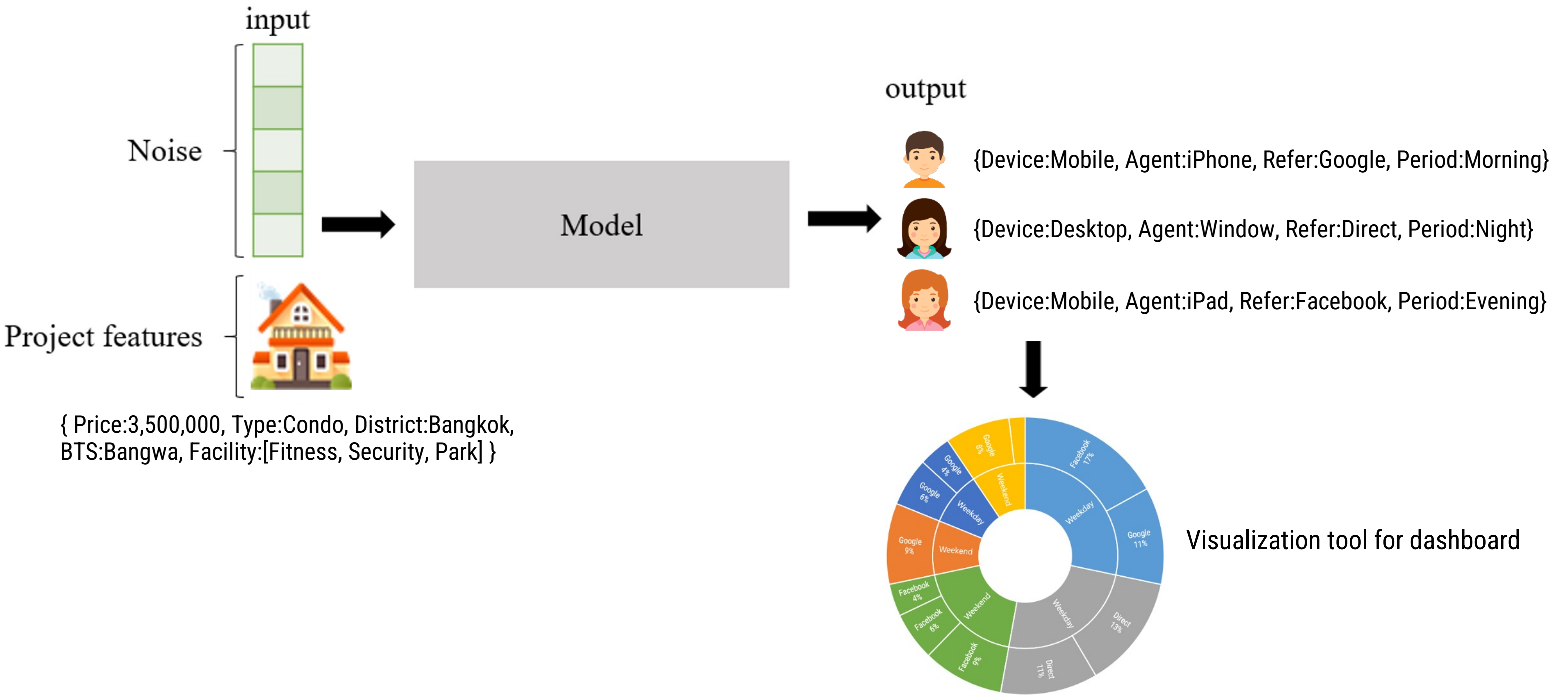
$$Y=f(x,c)$$

GAN can be conditioned (controlled) to generate things you want by concatenating additional information

Example of CGAN applications

Globally and Locally Consistent Image Completion [Iizuka et al., 2017]
StackGAN: Text to Photo-realistic Image Synthesis with Stacked GANs [Zhang et al. 2017]

Overview of our system

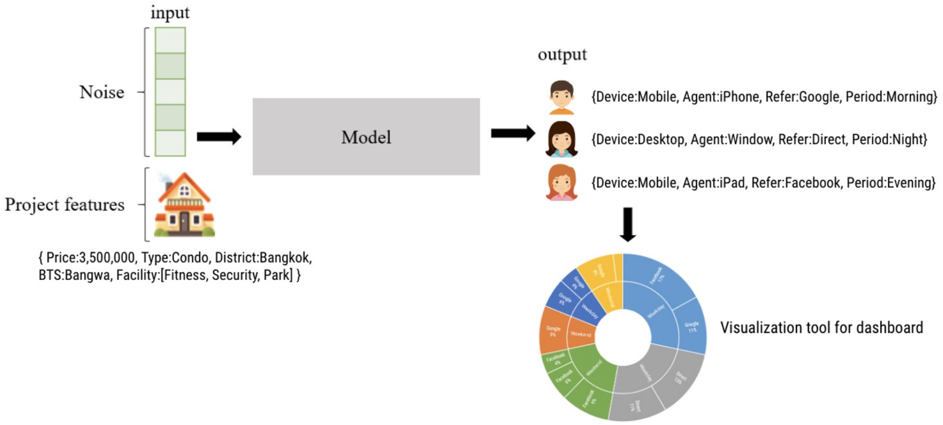


Experimental setup

~5000 projects, ~2 million log entries

- Held out 50 random projects as novel projects to generate
- Measure the distribution of generated logs vs real data

Average the performance over 10 runs

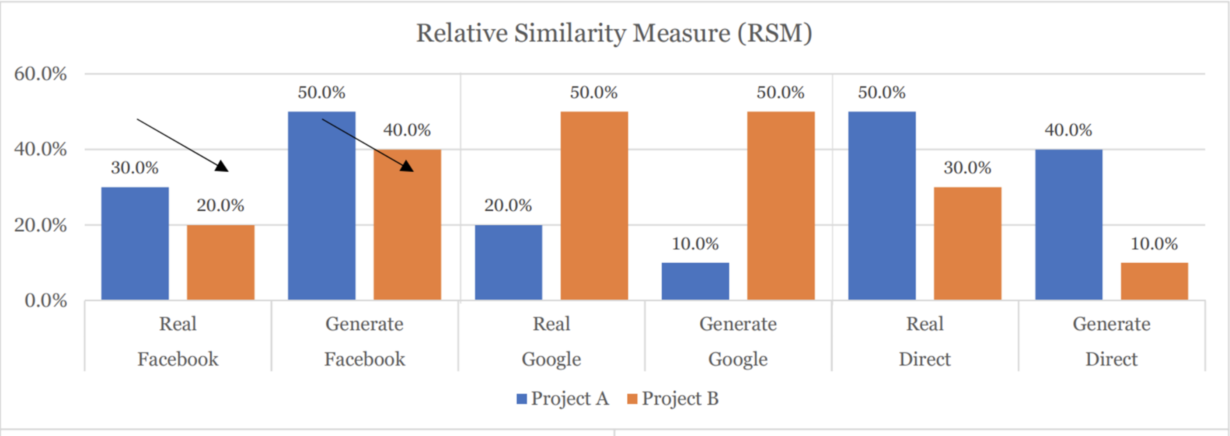


Metrics

RSM

Relative measure

Across project pairs



CHULA ENGINEERING
Foundation toward Innovation

Metrics

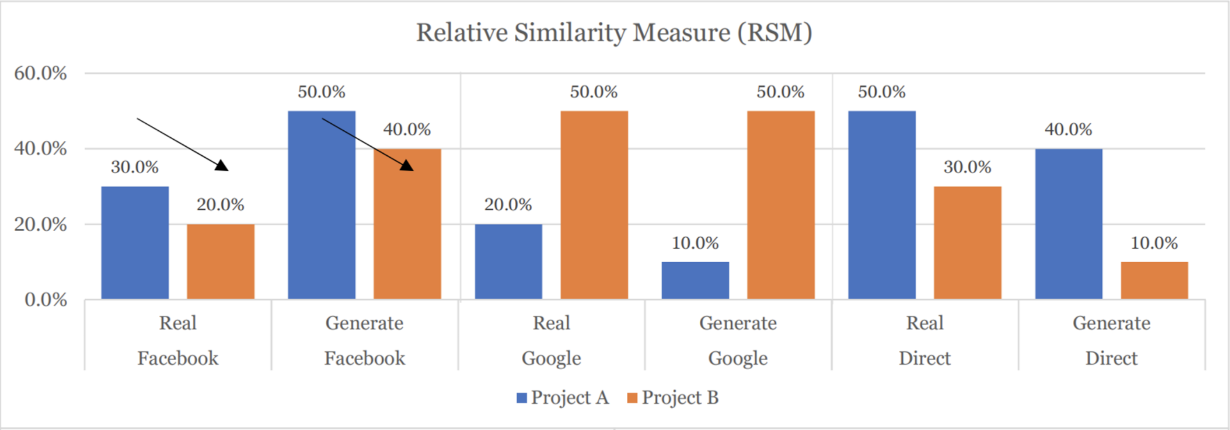
RSM

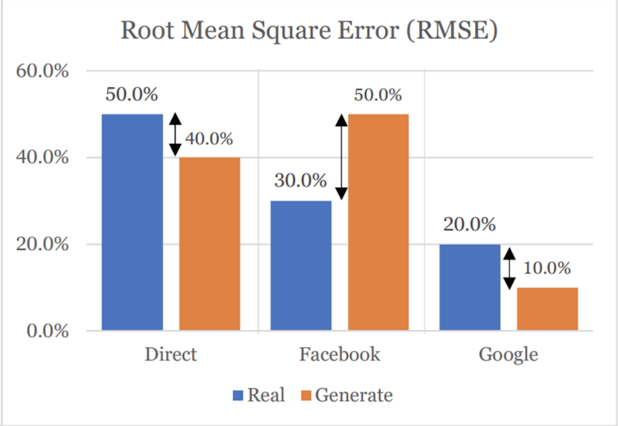
Relative measure

Across project pairs

RMSE

Absolute measure





Results

Model	RSM	RMSE
NN	54.7%	28.0%

Use the most similar project in the training data based on recommendation embeddings

Results

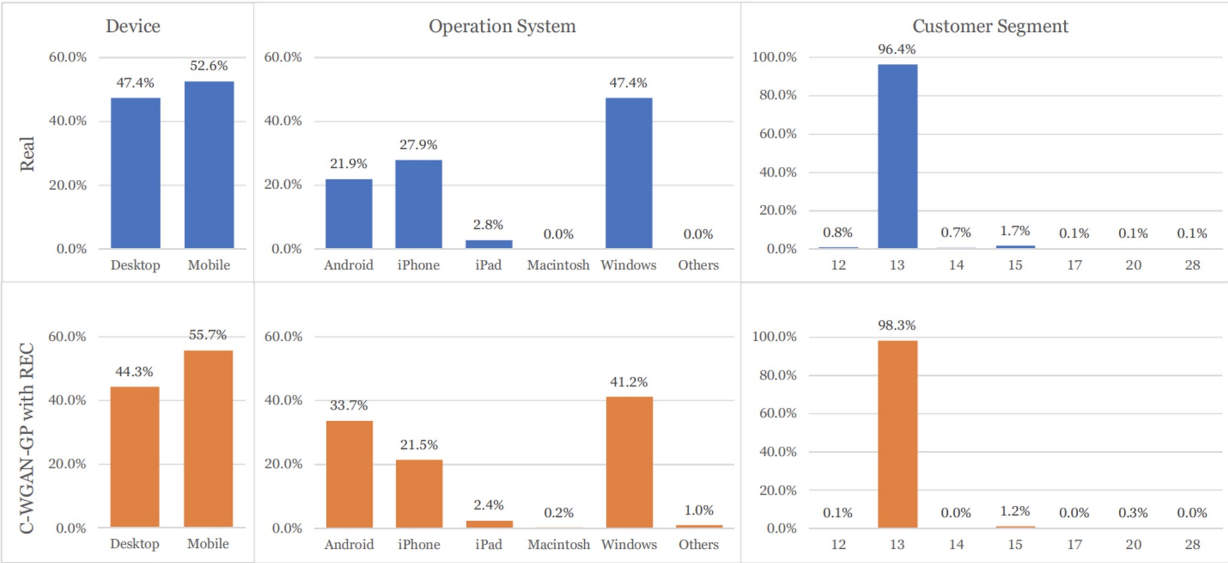
Model	RSM	RMSE
GAN	72.5%	16.2%
NN	54.7%	28.0%

Our model with recommender embedding

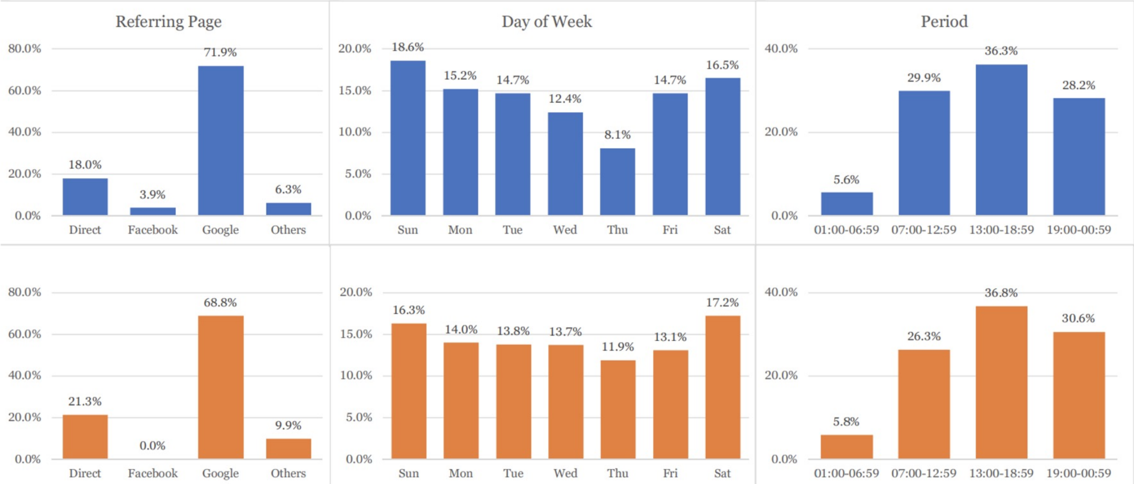
Results

Model	RSM	RMSE
GAN	72.5%	16.2%
VAE	65.3%	20.3%
NN	54.7%	28.0%

Instead of GAN use VAE



For more details
Generating Realistic Users Using Generative Adversarial Network With Recommendation-Based Embedding
https://ieeexplore.ieee.org/abstract/document/9016238



Text generation

- Generate text
  - Train a model to predicts words based on
    - Previous words

The cat is _____

สุขภาพ ไม่มีปัญหาอะไรรุนแรง แข็งแรงดี แต่ที่น่าห่วงอย่างเดียวคือ
ความเครียด คุณควรดูแลตัวเองให้ดี ไม่ควรเอาตัวเข้าไปเสี่ยงมากเกินไป
ถ้าเป็นไปได้ก็ไปพบแพทย์ ทำใจสบาย ๆ

Text generation

- Generate text
  - Train a model to predicts words based on
    - Previous words
    - Some conditioning (text, image, scenarios)

EX: OpenAI Codex https://openai.com/blog/openai-codex/

Input

← Use as input

```
class SiameseNetwork(nn.Module):
   # Todo: implement the network
   def __init__(self):


   def extract_feature(self, x):

   def forward(self, input1, input2):
```

Instructions

Create a model with resnet18, linear128, linear64

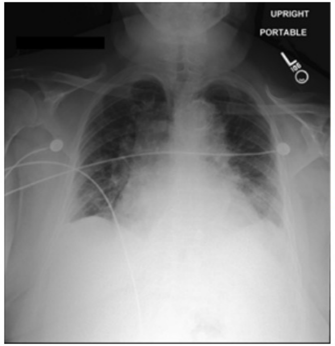
```
class SiameseNetwork(nn.Module):
    # Todo: implement the network
    def __init__(self):
        super(SiameseNetwork, self).__init__()
        self.resnet =
torchvision.models.resnet18(pretrained=True)
        self.linear1 = nn.Linear(512, 128)
        self.linear2 = nn.Linear(128, 64)

    def extract_feature(self, x):
        x = self.resnet(x)
        x = self.linear1(x)
        x = self.linear2(x)
        return x

    def forward(self, input1, input2):
        output1 = self.extract_feature(input1)
        output2 = self.extract_feature(input2)
        return output1, output2
```

Text generation

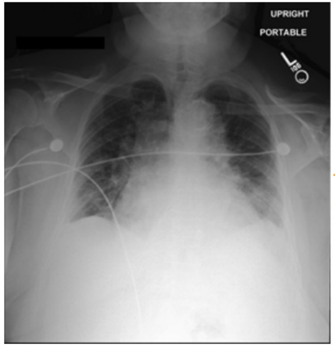
- Generate text
    - Train a model to predicts words based on
        - Previous words
        - Some conditioning (text, image, scenarios)



FINDINGS: In comparison to study performed on of ___ there is new mild pulmonary edema with small bilateral pleural effusions. Lung volumes have decreased with crowding of vasculature. No pneumothorax. Severe cardiomegaly is likely accentuated due to low lung volumes and patient positioning.

IMPRESSION: 1. New mild pulmonary edema with persistent small bilateral pleural effusions. 2. Severe cardiomegaly is likely accentuated due to low lung volumes and patient positioning.

X-ray description Model



→ Encoder (CNN) → Decoder (Transfomer) →

FINDINGS: In comparison to study per
mild pulmonary edema with small bilat
volumes have decreased with crowding
pneumothorax. Severe cardiomegaly is
lung volumes and patient positioning.

~100000 Xray images and reports used for training

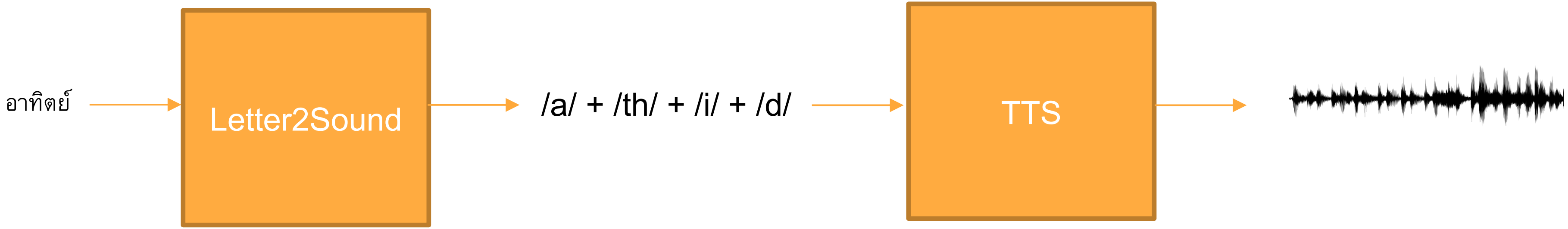
Example outputs

X-ray	Ground truth	LSP
	frontal and lateral views of the chest were obtained. there are streaky linear opacities at the lung bases which are likely due to atelectasis with chronic changes. no definite focal consolidation is seen. there is no pleural effusion or pneumothorax. no pneumothorax is seen. the aorta is calcified and tortuous. the cardiac silhouette is top normal to mildly enlarged. dual-lead left-sided pacemaker is seen with leads in the expected positions of the right atrium and right ventricle. chronic-appearing rib deformities on the right is again seen.	frontal and lateral views of the chest were obtained. there is a small left pleural effusion with overlying atelectasis. there is no focal consolidation, pleural effusion or pneumothorax. there is no pleural effusion or pneumothorax. the aorta is calcified and tortuous. the heart is mildly enlarged. a left-sided pacemaker is seen with leads in the expected position of the right atrium and right ventricle. the patient is status post median sternotomy and cabg. the lungs are otherwise clear.

For more detail: “Set Prediction in the Latent Space”
https://papers.nips.cc/paper/2021/hash/d61e9e58ae1058322bc169943b39f1d8-Abstract.html

Text-to-speech (TTS)

- Given a text input, output an utterance
  - Usually converts the text into sound units first (phonemes)
  - Recent research focus on emotion control



Human or Computer

A

B

1 out of 3 participants cannot identify the generated speech

Try our text2speech and speech2text demos at http://20.239.26.101/

For more detail: Spectral and Latent Speech Representation Distortion for TTS Evaluation.
https://www.isca-speech.org/archive/interspeech_2021/kongthaworn21_interspeech.html (Patent in process)

Conclusion

- Generative machine learning has come a long way
  - Could help increase productivity for many tasks
    - Human-in-the-loop research will be crucial
      - Using insights
      - Verifying results
      - Tweaking control parameters

Conclusion

- Generative machine learning has come a long way
  - Could help increase productivity for many tasks
    - Human-in-the-loop research will be crucial
  - Evaluating generative models is a challenge
    - task dependent, human evaluation not preferred
      - X-ray: extract labels from text and verify from the labels
      - Code: execution of the code
      - Voice: compare against pre-recorded using embedding features

Slides: https://bit.ly/aiiotgenerative

Conclusion

- Generative machine learning has come a long way
  - Could help increase productivity for many tasks
    - Human-in-the-loop research will be crucial
  - Evaluating generative models is a challenge
    - task dependent, human evaluation not preferred
  - Security concerns
    - Extensive research in detecting machine generated

PRO CYBER NEWS

Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case

Scams using artificial intelligence are a new challenge for companies

PHOTO: SIMON DAWSON/BLOOMBERG NEWS

By Catherine Stupp

Updated Aug. 30, 2019 12:52 pm ET

MOST POPULAR NEWS

1. Where State Abortion Laws Stand if Roe v. Wade Is Overturned
2. They Kept Paying When Student Loan Debt Paused
3. NFT Sales Are Flatlining
4. Cerebral's Preferred Pharmacy Truepill Halts Adderall Prescriptions for All Customers
5. Immigrants to Get Extension for Expiring or Expired U.S. Work

https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402