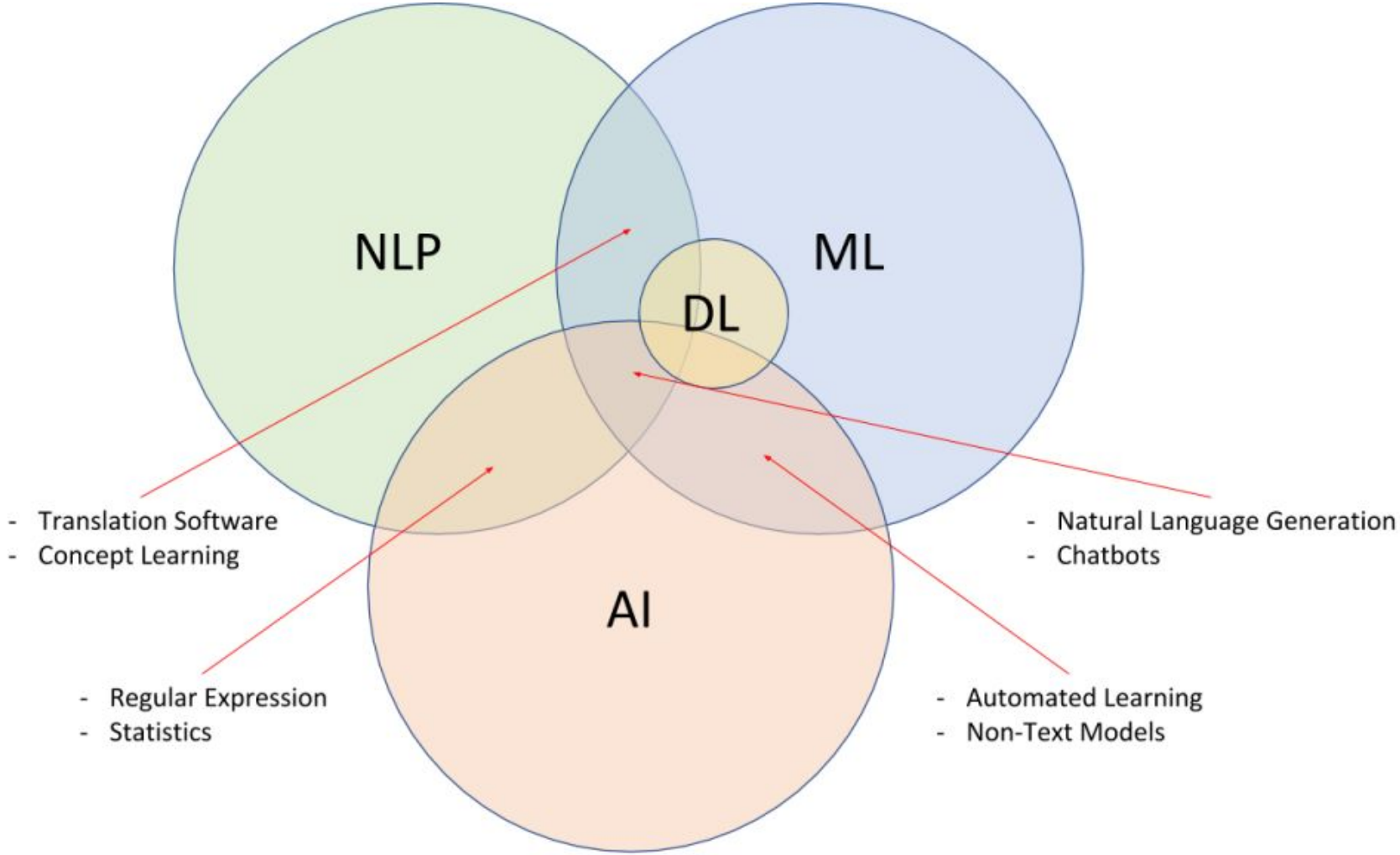
# MIDS W207
# Applied Machine Learning

Week 9
Live Session Slides
Fall 2023


NLP
ML
DL
AI
- Translation Software
- Concept Learning
- Natural Language Generation
- Chatbots
- Regular Expression
- Statistics
- Automated Learning
- Non-Text Models

# Sentiment Analysis

My experience so far has been fantastic!

POSITIVE

The product is ok I guess

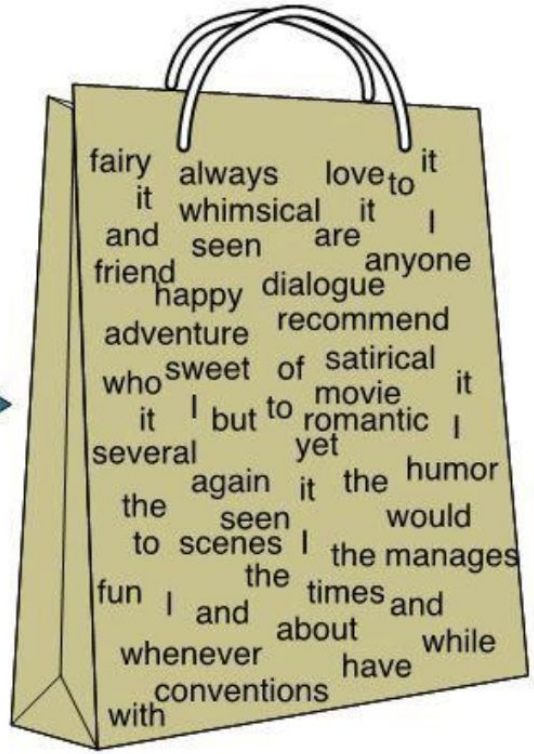
NEUTRAL

Your support team is useless

NEGATIVE

# The Bag of Words Representation

I love this movie! It's sweet, but with satirical humor. The dialogue is great and the adventure scenes are fun... It manages to be whimsical and romantic while laughing at the conventions of the fairy tale genre. I would recommend it to just about anyone. I've seen it several times, and I'm always happy to see it again whenever I have a friend who hasn't seen it yet!




| | |
|---|---|
| it | 6 |
| I | 5 |
| the | 4 |
| to | 3 |
| and | 3 |
| seen | 2 |
| yet | 1 |
| would | 1 |
| whimsical | 1 |
| times | 1 |
| sweet | 1 |
| satirical | 1 |
| adventure | 1 |
| genre | 1 |
| fairy | 1 |
| humor | 1 |
| have | 1 |
| great | 1 |
| … | … |

Sentence 1 : He is a good doctor
Sentence 2: She is a good scientist
Sentence 3: Doctor and Scientist are good

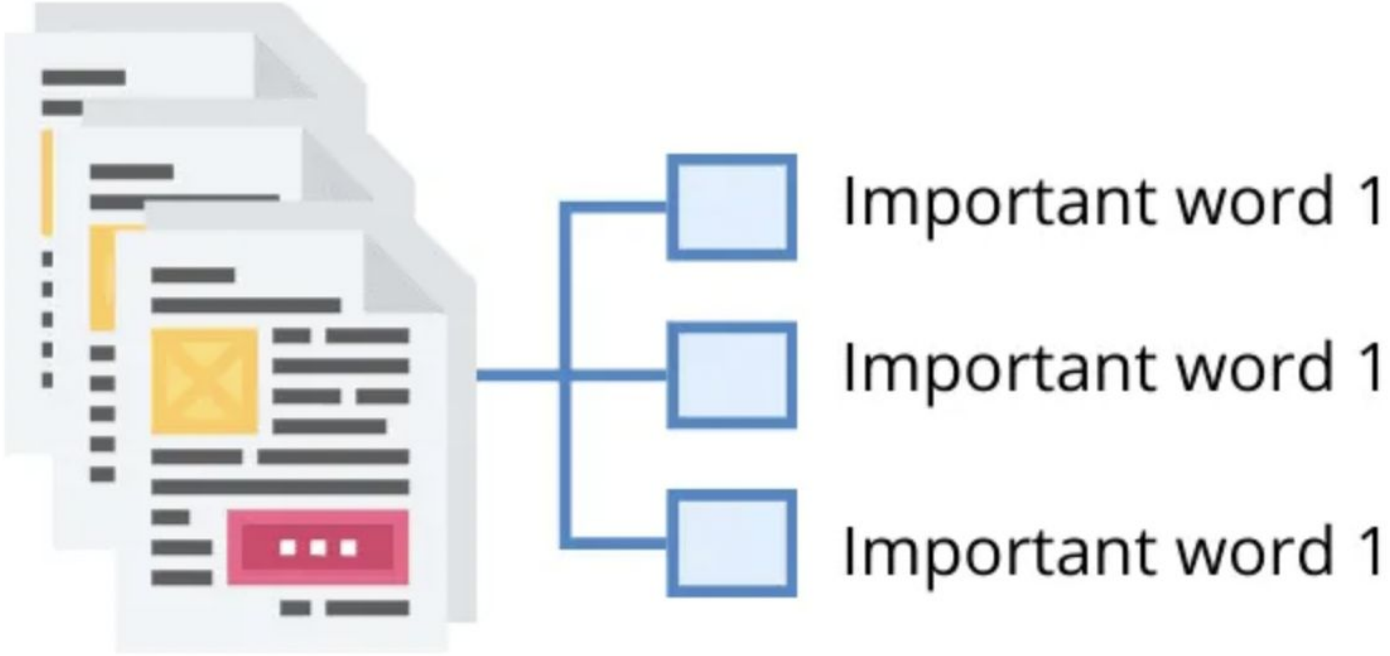
Sentence 1 : good doctor
Sentence 2: good scientist
Sentence 3: doctor scientist good

| Words | Frequency |
|---|---|
| good | 3 |
| doctor | 2 |
| scientist | 2 |

vectors

| | f1 | f2 | f3 | |
|---|---|---|---|---|
| | good | doctor | scientist | output |
| Sentence 1 | 1 | 1 | 0 | |
| Sentence 2 | 1 | 0 | 1 | |
| Sentence 3 | 1 | 1 | 1 | |

# TF - IDF



Sentence 1 : good doctor
Sentence 2: good scientist
Sentence 3: doctor scientist good

**TF * IDF**

Term Frequency= No. of repetitive words in a sentence/ No. of words in a sentence

Inverse Document Frequency= Log (No. of sentences/ No of sentences containing words)

| Words | Frequency |
|---|---|
| good | 3 |
| doctor | 2 |
| scientist | 2 |

**TF**

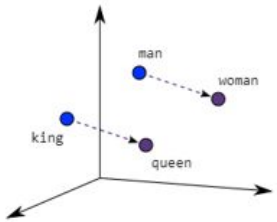
| | S1 | S2 | S3 |
|---|---|---|---|
| good | 1/2 | 1/2 | 1/3 |
| doctor | 1/2 | 0 | 1/3 |
| scientist | 0 | 1/2 | 1/3 |

**IDF**

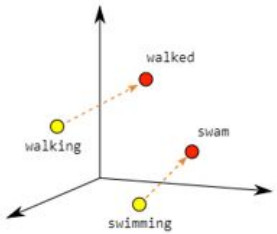
| Words | IDF |
|---|---|
| good | Log (3/3)=0 |
| doctor | Log (3/2) |
| scientist | Log (3/2) |

| | f1 good | f2 doctor | f3 scientist | output |
|---|---|---|---|---|
| S1 | 0 | ½ * log (3/2) | 0 | |
| S2 | 0 | 0 | ½ * log (3/2) | |
| S3 | 0 | ⅓ * log (3/2) | ⅓ * log (3/2) | |

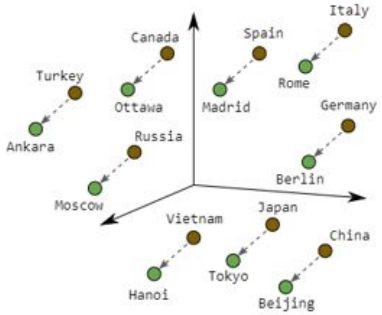
# Word2Vec




Male-Female

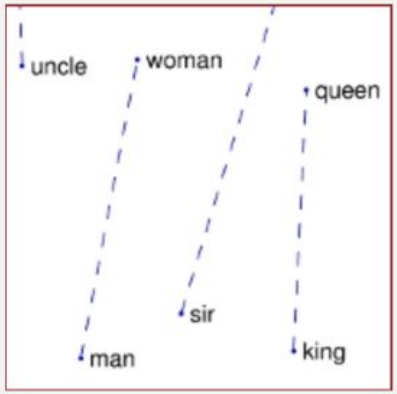



Verb Tense

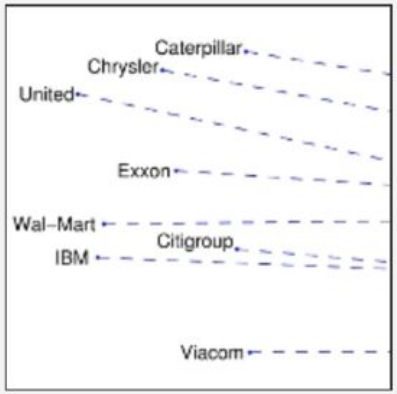



Country-Capital

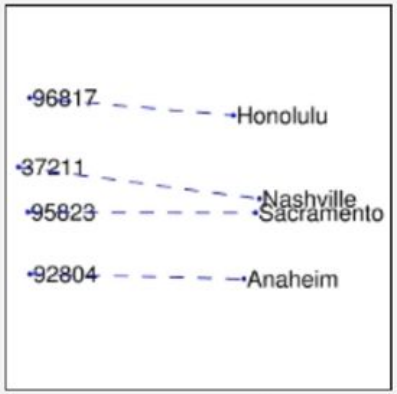
# GloVe



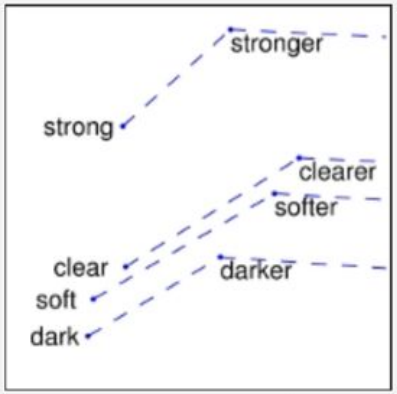

man - woman




company - ceo

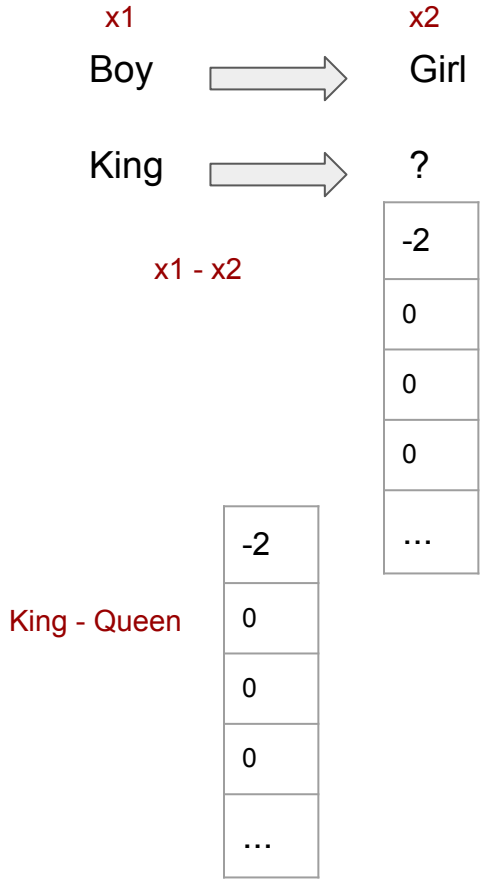



city - zip code

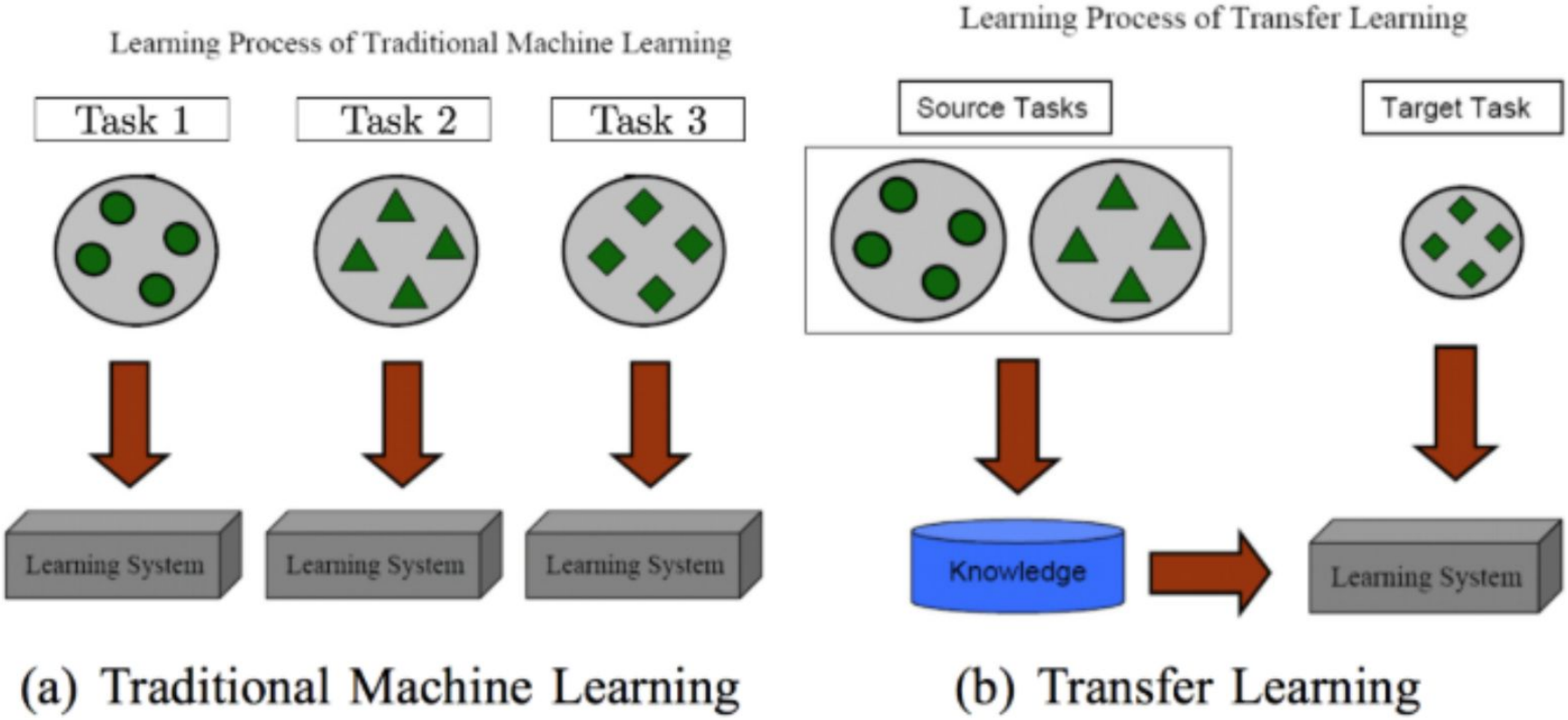


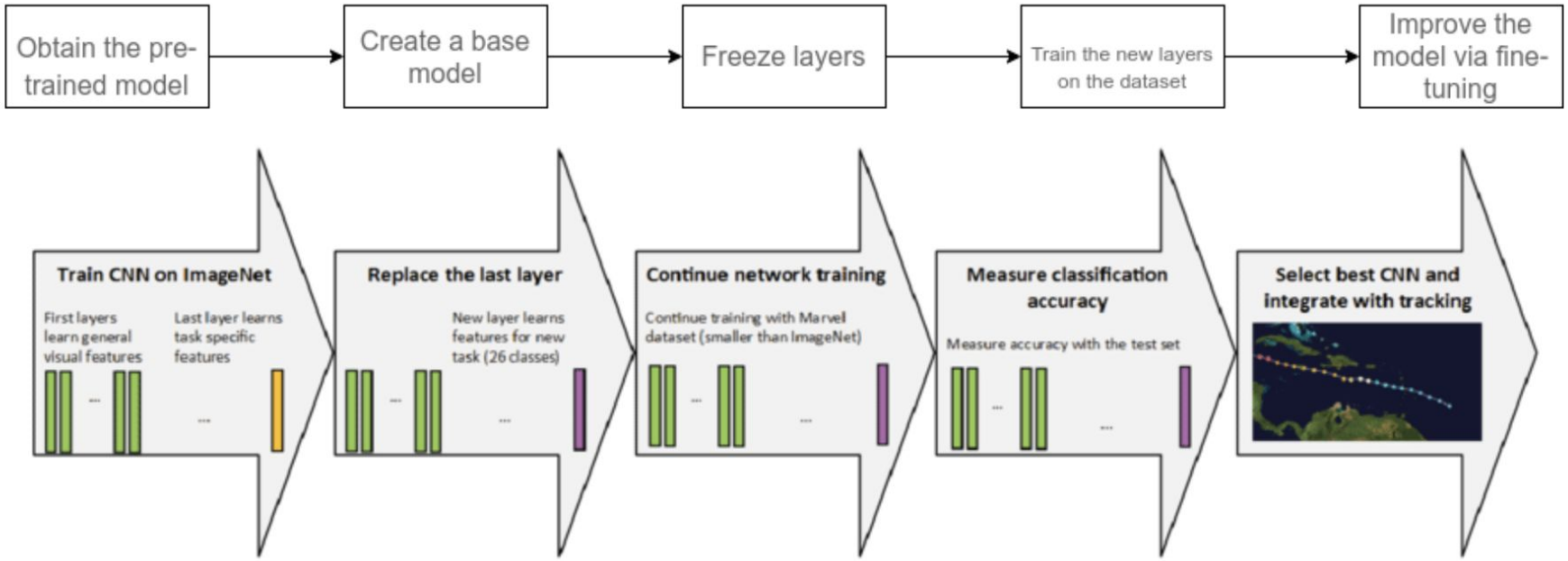

comparative - superlative

| | Boy | Girl | King | Queen | Apple | Mango |
|---|---|---|---|---|---|---|
| Gender | -1 | 1 | -0.92 | 0.93 | 0 | 0.1 |
| Royal | 0.01 | 0.02 | 0.95 | 0.96 | -0.02 | 0.01 |
| Age | 0.03 | 0.02 | 0.7 | 0.6 | 0.95 | 0.92 |
| Food | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| 300 | | | | | | |




| | x1 Boy | x2 Girl | King | Queen | Apple | Mango |
|---|---|---|---|---|---|---|
| Gender | -1 | 1 | -0.92 | 0.93 | 0 | 0.1 |
| Royal | 0.01 | 0.02 | 0.95 | 0.96 | -0.02 | 0.01 |
| Age | 0.03 | 0.02 | 0.7 | 0.6 | 0.95 | 0.92 |
| Food | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| 300 | | | | | | |


Learning Process of Traditional Machine Learning
Task 1
Task 2
Task 3
Learning System
Learning System
Learning System
(a) Traditional Machine Learning
Learning Process of Transfer Learning
Source Tasks
Target Task
Knowledge
Learning System
(b) Transfer Learning


Obtain the pre-trained model
Create a base model
Freeze layers
Train the new layers on the dataset
Improve the model via fine-tuning
Train CNN on ImageNet
First layers learn general visual features
Last layer learns task specific features
Replace the last layer
New layer learns features for new task (26 classes)
Continue network training
Continue training with Marvel dataset (smaller than ImageNet)
Measure classification accuracy
Measure accuracy with the test set
Select best CNN and integrate with tracking

# Code Review