Project 158
செயற்றிட்டம் 158



DP EDUCATION Coding School



AI and Machine



LIVEWORKSHEETS

- Machine learning என்பதை இலகுவாக எடுப்போமானால் ஏதாவது ஒரு machine இற்கு இறந்த கால data மூலம் கற்பித்து எதிர்காலத்தில் ஏதாவது தீர்மானம் எடுக்கும் வாய்ப்பை அந்த machine இற்கு பெற்றுக் கொடுப்பதாகும்.
- Machine learning இன் பிரதான பகுதிகள் மூன்று.
    - Supervised Learning
    - Unsupervised Learning
    - Reinforcement Learning
- இங்கு நாங்கள் Supervised Learning பற்றி பார்ப்போம்.

Supervised Learning

- Supervised Learning களில் label பண்ணப்பட்ட data மூலம் machine train பண்ணப்படும்.
- உங்களுக்கு பூனை ஒன்றின் image ஐ காட்டி அது பூனை ஒன்று என குறிப்பிடும் போது அது label பண்ணப்பட்ட data ஒன்றாகும்.
- Label பண்ணப்பட்ட data என்பதை கீழே உள்ள உதாரணத்தின் மூலம் விளங்கிக் கொள்வோம்.
- இங்கு நாம் machine learning model ஒன்றை training பண்ணுவோம். exam ஒன்றின் புள்ளிகளுக்கு ஏற்ப மாணவர்கள் தேர்ச்சி அடைந்தார்களா இல்லையா என்பதை தீர்மானிக்க.
- அதற்காக கீழ் உள்ள data set ஐ கவனித்துப் பார்ப்போம்.

LIVEWORKSHEETS

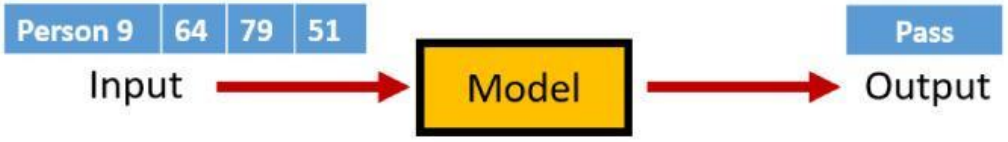
Person 1	80	60	65	**Pass**
Person 2	70	54	63	**Pass**
Person 3	61	47	86	**Fail**
Person 4	50	50	83	**Pass**
Person 5	74	31	66	**Fail**
Person 6	58	81	43	**Fail**

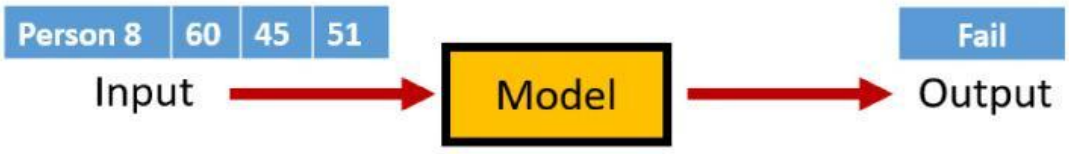
- இங்கு மூன்று exam களில் பெற்றுக்கொண்ட புள்ளிகளுக்கு ஏற்ப அந்த நபர் தேர்ச்சி பெற்றாரா இல்லையா என்பது Result எனும் நிரலில் காட்டப்படும்.
- இங்குள்ளது label பண்ணப்பட்ட data set ஆகும்.
- இங்கு நாம் input ஐ பெற்று கொடுக்கும் data set அதாவது ஒவ்வொருவரும் exam களில் பெற்றுக்கொண்ட புள்ளிகள், மேலும் அந்த புள்ளிகளுக்கு ஏற்ப அவர்களுக்கு கிடைத்த result ஒன்று இருக்கும். அது output ஆகும்.
- இந்த data மூலம் முதலில் machine learning modal ஒன்றை training பண்ண வேண்டும்.
- அப்போது இந்த data களுக்கு ஏற்ப exam pass பண்ண எவ்வகையான புள்ளிகள் இருக்க வேண்டும், exam ஒன்று fail பண்ண எவ்வகையான புள்ளிகள் இருக்க வேண்டும்

LIVEWORKSHEETS

போன்ற pattern ஒன்றை machine இற்கு உருவாக்கிக் கொள்ள முடியும்.

- அந்த pattern தான் machine learning modal ஒன்றாக machine மூலம் store பண்ணிக்கொள்ளப்படும்.
- "இவ்வாறு machine learning modal மாதிரியைப் பயிற்சி செய்த பிறகு, மாணவர்கள் பெற்ற மதிப்பெண்களை இந்த மாதிரியில் input ஆக கொடுத்து, அவர்கள் தேர்ச்சி அடைவார்களா இல்லையா என்பதை machine எவ்வாறு கணிக்கிறது என்பதை நாம் பார்க்கலாம்."





LIVEWORKSHEETS

- ❖ சில சமயங்களில் அவர்கள் பெற்றுத் தரும் விடை பிழையாக முடியும்.
- ❖ எங்களுக்கு இவ்வாறு கிடைக்கும் result இன் accuracy ஐ கூட்டுவதற்கு எங்கள் machine learning modal ஐ training பண்ணுவதற்கு எடுக்கும் data set இல் உள்ள data இன் அளவை கூட்ட வேண்டும்.
- ❖ Supervised learning use பண்ணும் Algorithms பின்வருமாறு அமையும்.
    - Linear Regression
    - Logistic Regression
    - Support Vector machine (SVM)
    - K-Nearest Neighbors (KNN)

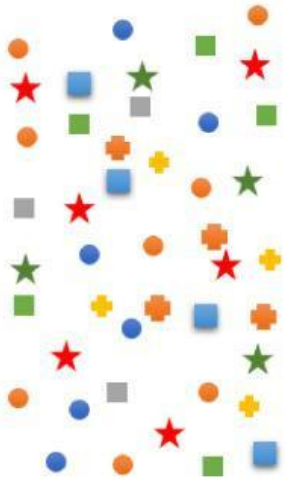
- ❖ Supervised learning களில் அறிந்து கொள்ள வேண்டிய பிரதான குறிப்பு machine learning modal ஐ training பண்ணுவது label பண்ணப்பட்ட data ஐ பயன்படுத்தியாகும்.

Unsupervised Learning

- ❖ Unsupervised learning களில் machine learning modal ஐ training பண்ணுவதற்கு label பண்ணப்படாத data set ஒன்று பயன்படுத்தப்படும்.

LIVEWORKSHEETS

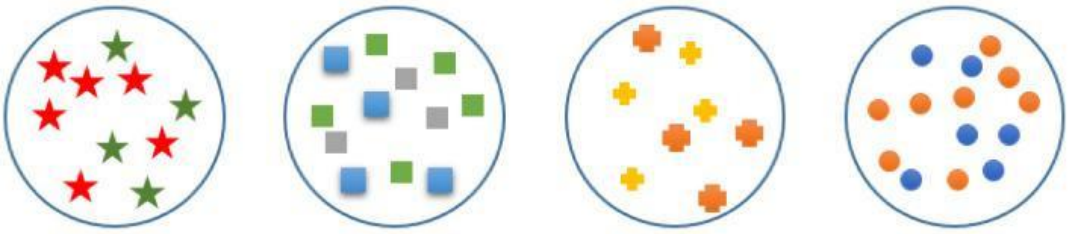
- இங்கு நாம் data மூலம் machine learning modal ஐ training பண்ணினாலும் output ஐ machine இற்கு அறிமுகப் படுத்துவதில்லை.
- பின்வருமாறு data set உண்டென்று நினைப்போம்.



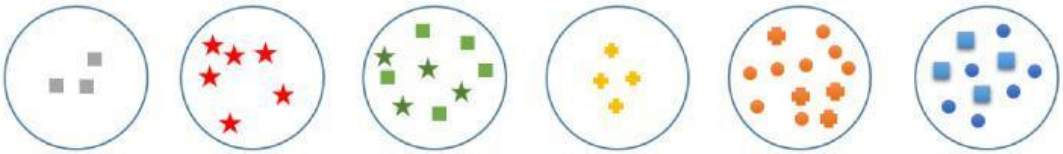
- மேலுள்ளவாறு data set ஒன்றை நாம் Supervised learning களில் பயன்படுத்தியிருந்தால் அதில் உள்ள வடிவங்களை அறிமுகப்படுத்த வேண்டி இருக்கும். அதாவது இங்கு உள்ள சிறியது எது, வட்டம் எது, நட்சத்திரம் எது என்று நாம் label பண்ணி அறிமுகப்படுத்த வேண்டியிருக்கும்.
- ஆனால் Unsupervised learning களில் நாம் இந்த data label பண்ணுவது இல்லை.

LIVEWORKSHEETS

- Unsupervised learning களின் machine மூலம் data set வகைபடுத்தப்படும். அதற்கு machine இற்கு புரியும் Pattern பயன்படுத்தப்படும்.
- ஆனால் இங்கு வகைப்படுத்தப்பட்டாலும் இவற்றுக்கு கூறப்படும் பெயர் machine இற்கு தெரியாது.
- இங்கு வடிவங்களுக்கு ஏற்ப பின்வருமாறு வகைப்படுத்தப்படும்.

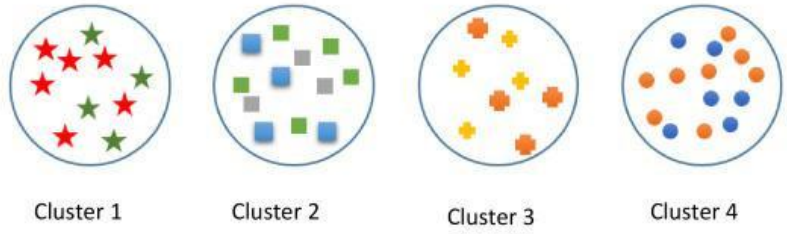


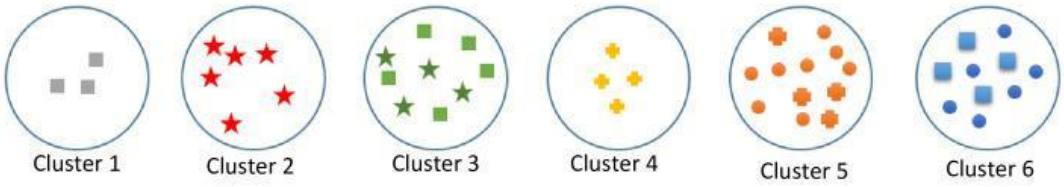
- மேலும் அந்த Algorithms இற்கு நிறத்தை அடையாளம் கண்டு கொள்ள முடியுமாயின் பின்வருமாறு வடிவங்களை வேறாக்க முடியும்.



LIVEWORKSHEETS

❖ Unsupervised learning களின் இந்த data set ஐ ஒவ்வொரு கூட்டத்திற்கு வேறாக்கி கொள்வதை நாம் clustering என்று அறிமுகப்படுத்துவோம்.





❖ அவ்வாறு எங்களுக்கு பயன்படுத்த முடியுமான clustering முறைகள் பல உள்ளன.அந்த Algorithms பின்வருமாறு அமையும்.

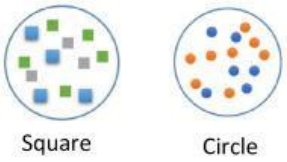
- K-Means Clustering
- Hierarchical Clustering
- Principal Component Analysis (PCA)
- Singular value Decomposition

Summary

LIVEWORKSHEETS

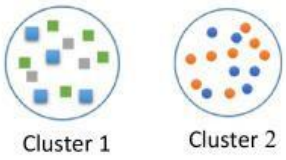
Unsupervised learning

- Machine learning modal training பண்ணுவதற்கு Label பண்ணப்பட்ட data set ஒன்று பயன்படுத்தப்படும்.
- Output



Supervised learning

- Machine learning modal trainingபண்ணுவதற்கு Label பண்ணப்படாத data set ஒன்று பயன்படுத்தப்படும்.
- Output



LIVEWORKSHEETS