

The background features a series of concentric, light gray circles centered on the slide. Overlaid on these are stylized, light blue circuit-like lines with small circles at the ends, resembling data paths or network connections. These lines are positioned in the corners of the slide.

DATA MODELLING

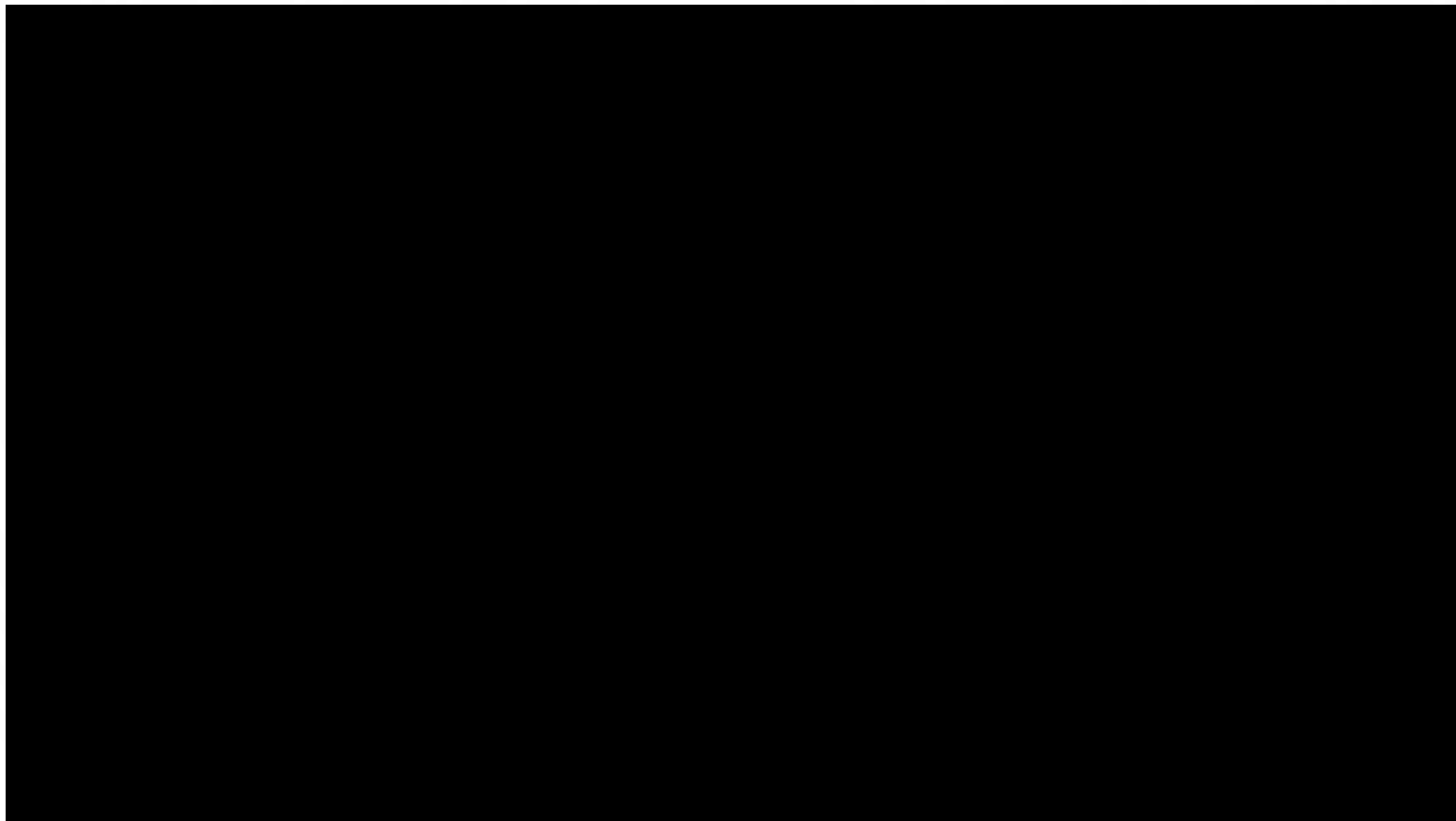
JENS BAETENS

DEFINITIE

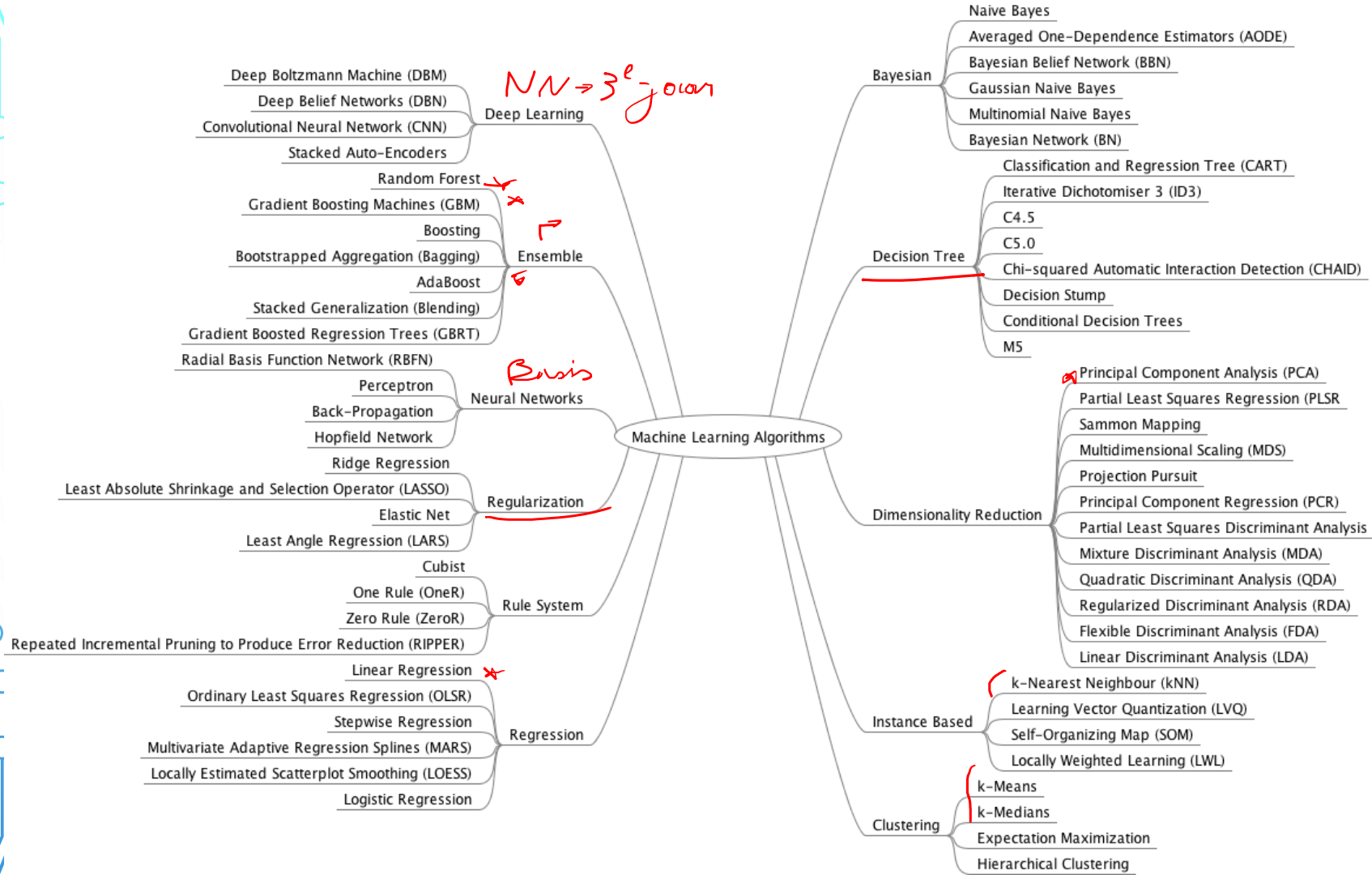
Opbouwen van een model om waarde en inzichten uit de data te halen en een antwoord te bieden op de gewenste vraag.

Stappen:

- 1 - Kies een of meerdere ML technieken om te gebruiken
- Trainen van het model op basis van de data
- Evalueren van het model op nieuwe data



https://www.youtube.com/watch?v=DjldgkCWUcA&list=RDCMUCoSiBUY5b9ixwuKfS99_rSw&index=1



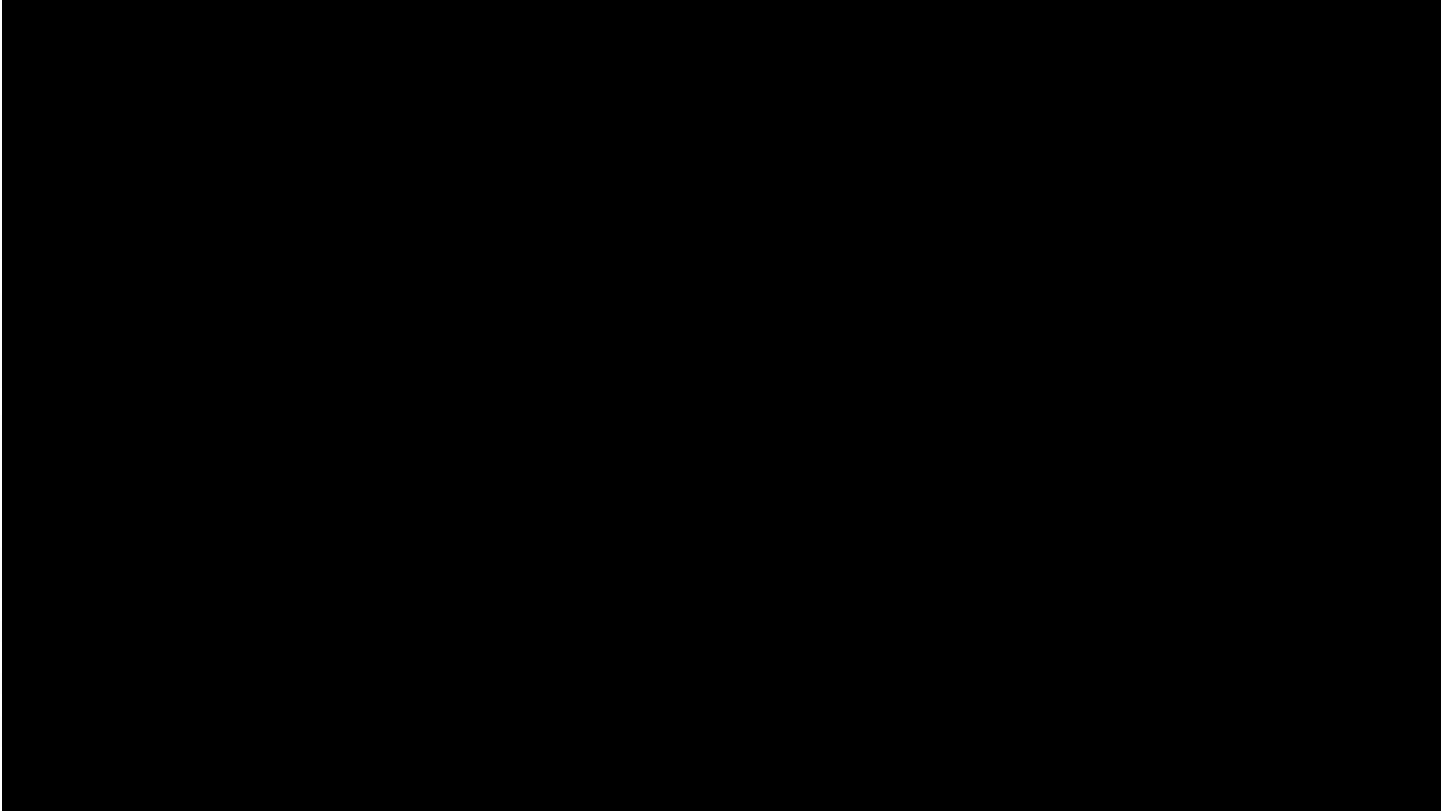
SOORTEN TECHNIEKEN

Gebruik van data en labels -> Supervised Machine Learning

Gebruik van data -> Unsupervised Machine Learning

Gebruik van regels -> Reinforcement Learning

SUPERVISED VS UNSUPERVISED



https://www.youtube.com/watch?v=cfj6yaYE86U&list=RDCMUCoSiBUY5b9ixwuKfS99_rSw&index=4

SUPERVISED LEARNING

met we voorzien
target

Attributes / features / kolommen

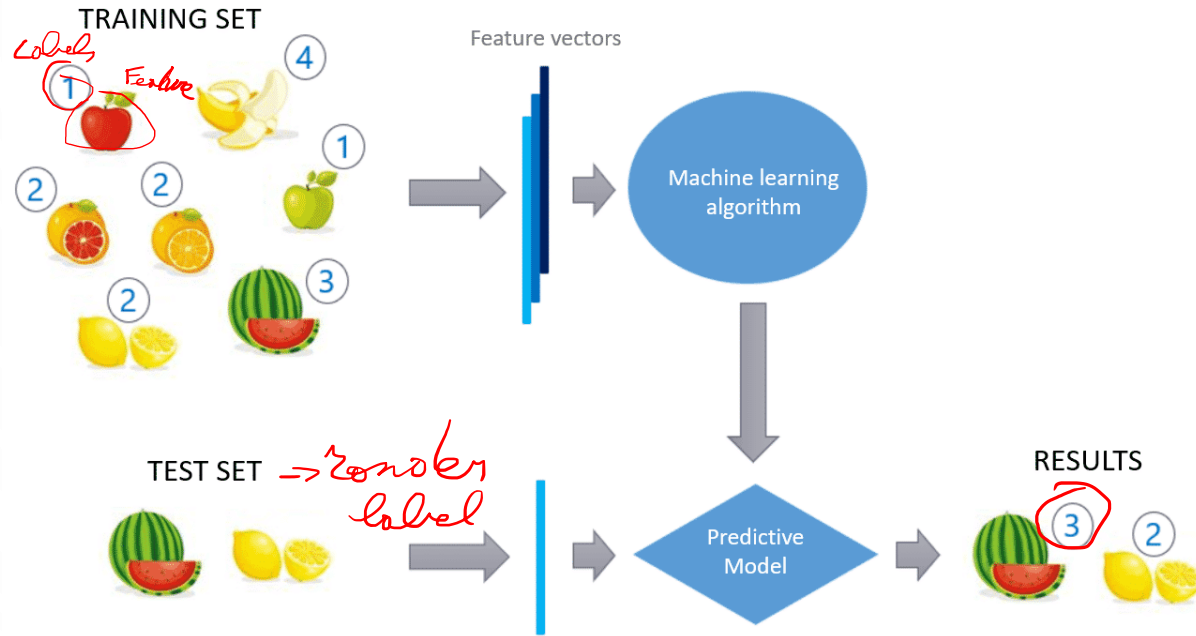
Inputs

Labels

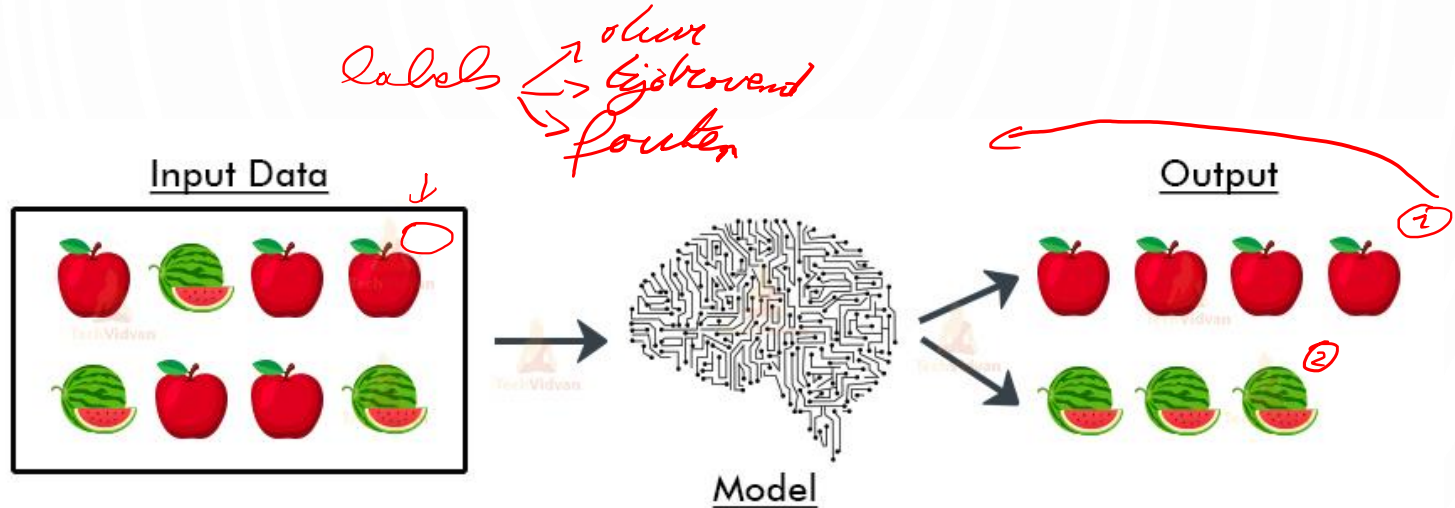
| # Avg. Area I... | # Avg. Area ... | # Avg. Area ... | # Avg. Area ... | # Area Popul... | # Price | ▲ Address |
|--------------------|--------------------|-------------------|-----------------|--------------------|--------------------|---|
| 79545.45857431678 | 5.682861321615587 | 7.009188142792237 | 4.09 | 23086.800502686456 | 1059033.5578701235 | 208 Michael Ferry Apt. 674 Laurabury, NE 37010-5101 |
| 79248.64245482568 | 6.0028998082752425 | 6.730821019094919 | 3.09 | 40173.07217364482 | 1505890.91484695 | 188 Johnson Views Suite 079 Lake Kathleen, CA 48958 |
| 61287.067178656784 | 5.865889840310001 | 8.512727430375099 | 5.13 | 36882.15939970458 | 1058987.9878760849 | 9127 Elizabeth Stravenue Danieltown, WI 06482-3489 |
| 63345.24004622798 | 7.1882360945186425 | 5.586728664827653 | 3.26 | 34310.24283090706 | 1260616.8066294468 | USS Barnett FPO AP 44820 |
| 59982.197225708034 | 5.040554523106283 | 7.839387785120487 | 4.23 | 26354.109472103148 | 630943.4893385402 | USNS Raymond FPO AE 09386 |

rijke observaties training exampel

SUPERVISED LEARNING



UNSUPERVISED LEARNING



ook gebruikt voor supervised



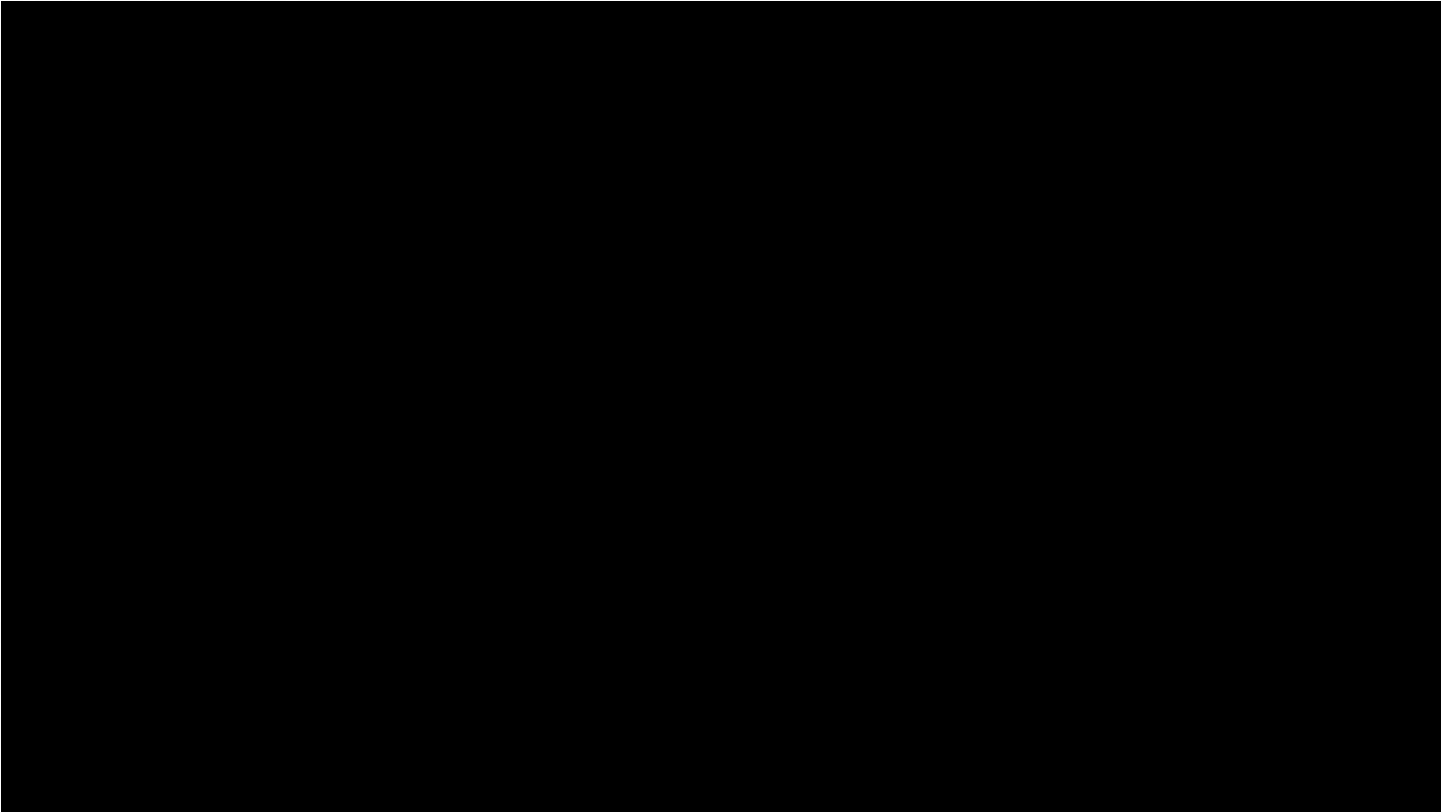
REINFORCEMENT LEARNING

Omvat technieken om modellen te bouwen die leren wat de beste actie om een zo hoog mogelijke score te bekomen.

Geen data gebruikt maar wel de regels van de omgeving waarin het moet werken

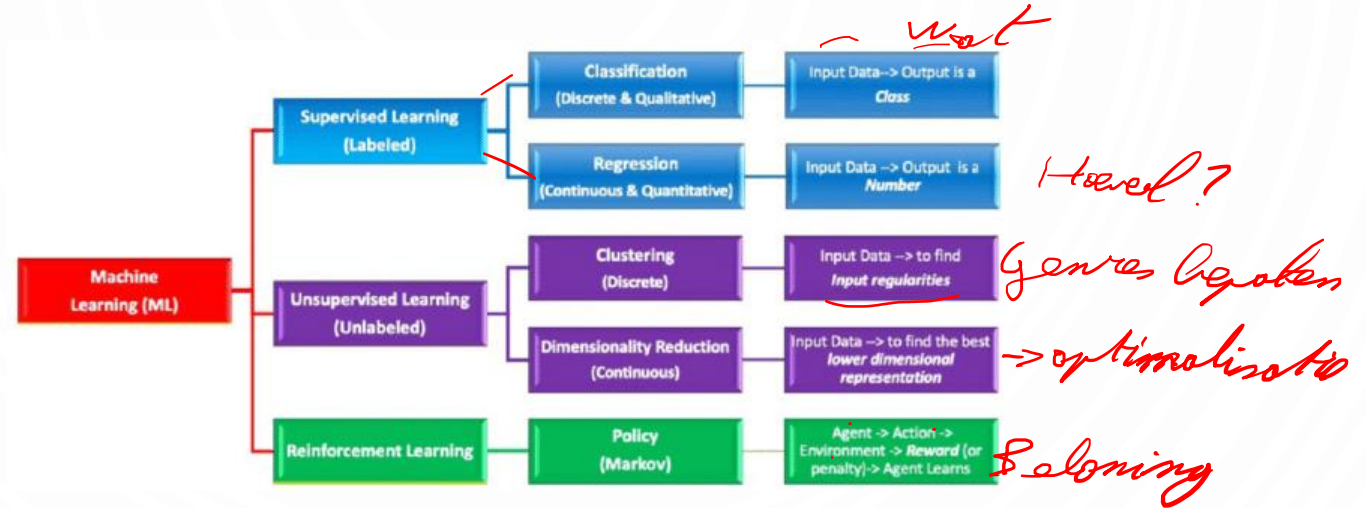


REINFORCEMENT LEARNING



<https://www.youtube.com/watch?v=Lu56xVIZ40M>

TECHNIEKEN



SOORTEN TECHNIEKEN - DOEL

Hoeveel gaat het zijn? -> Regression techniques

Is het type A of B? -> Classification techniques

Zijn deze gelijk? -> Clustering techniques

Is het vreemd? -> Anomaly detection techniques

Wat moet ik doen? -> Recommendation techniques

