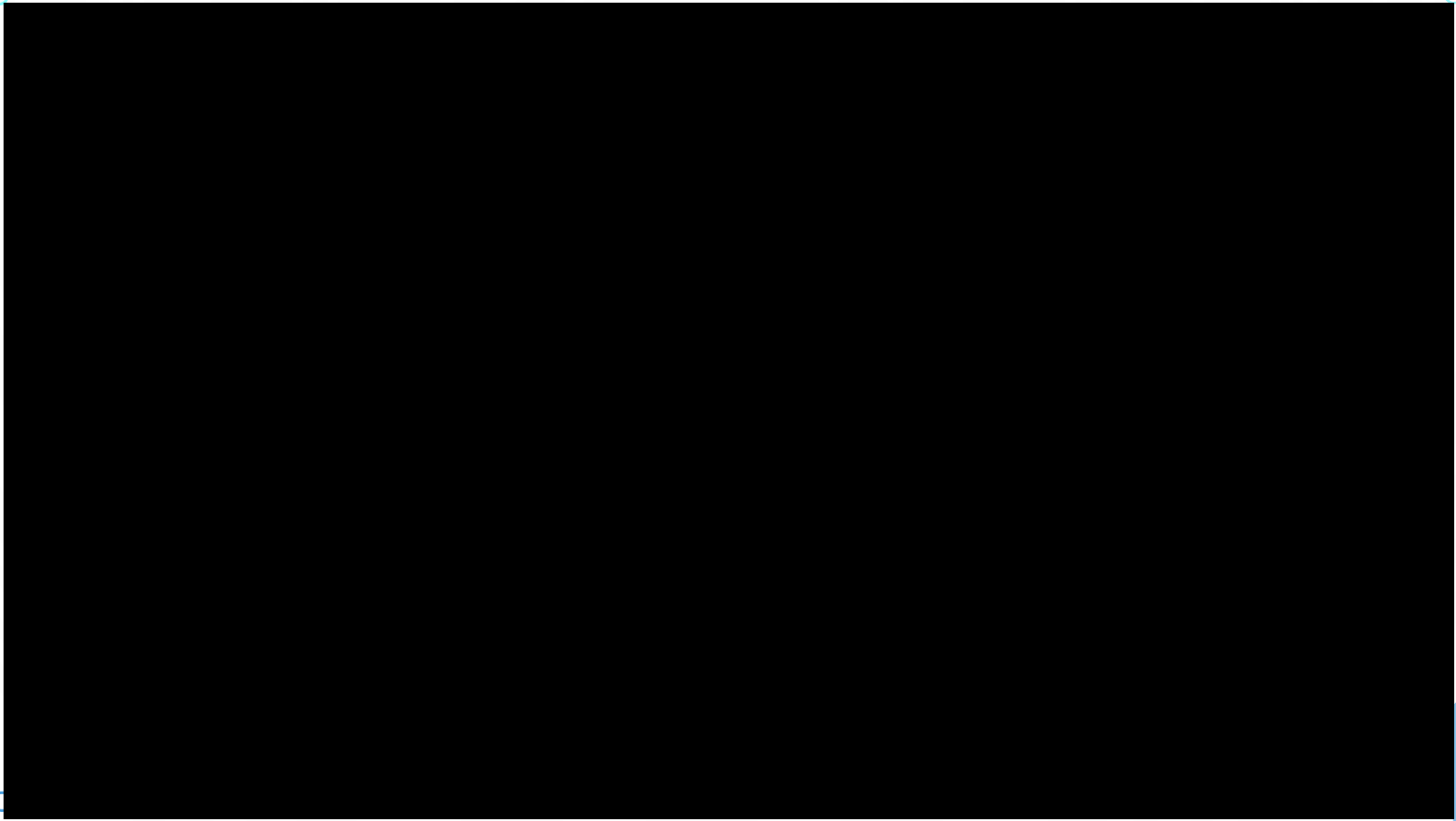


The background features a series of concentric circles centered on the slide, creating a ripple effect. In the corners, there are stylized circuit board traces in light blue and teal, with small circles at the end of the lines, resembling nodes or components.

ENSEMBLE LEARNING

JENS BAETENS





<https://www.youtube.com/watch?v=Un9zObFjBH0>



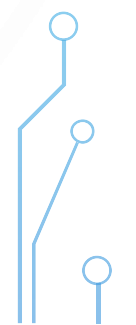
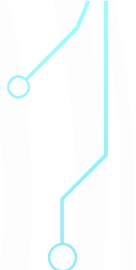

KENMERKEN

Meerdere modellen die samen de correcte output proberen te voorspelen

Gaat zowel voor classificatie als regressie

Meerdere technieken kunnen gecombineerd worden

=> Doel is om een betere voorspelling te bekomen dan de afzonderlijke modellen



STACKING



Decision making model: Majority voting of een nieuwe classifier

Kan op basis van de predicties maar ook op basis van de kansen van elke classifier

BAGGING – BOOTSTRAP AGGREGATING

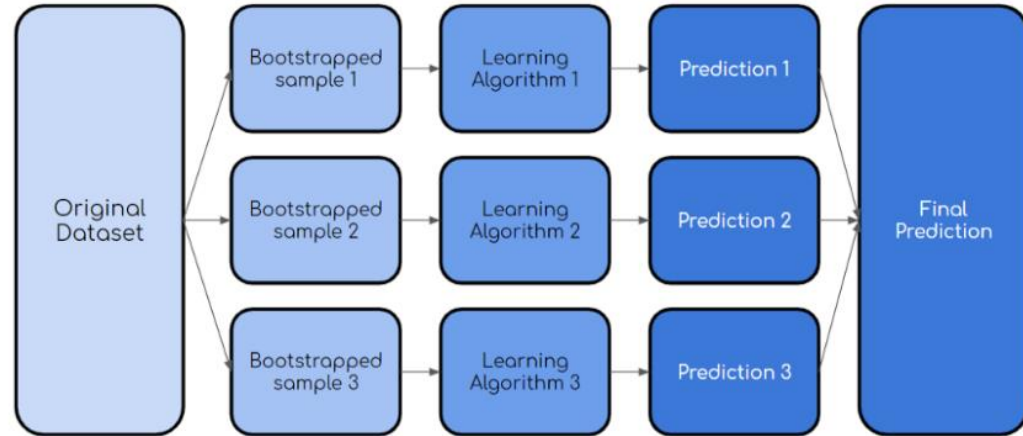
Meerdere modellen van hetzelfde algoritme

Train elk model met deel van de data
(Vaak rond 60%)

Samplen gebeurt met teruglegging

Standaard majority voting

Verminderd overfitting



<https://towardsdatascience.com/ensemble-learning-bagging-and-boosting-explained-in-3-minutes-2e6d2240ae21>

BOOSTING

Gelijkaardig aan bagging

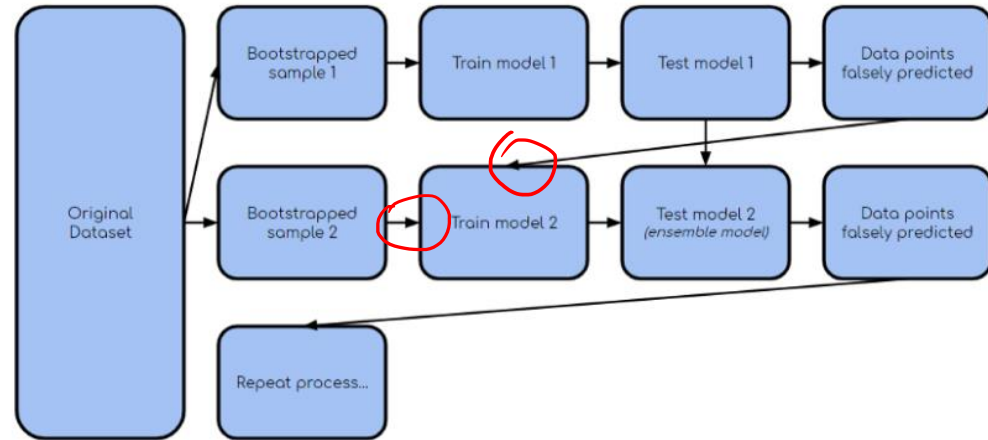
Maar training moet sequentieel gebeuren

Misgeclassificeerde samples grotere kans om gebruikt te worden om het volgende model te trainen

Meest gekende algoritme: AdaBoosting

VB: Face detection in camera's

predictie kan parallel



<https://towardsdatascience.com/ensemble-learning-bagging-and-boosting-explained-in-3-minutes-2e6d2240ae21>

XGBOOST

<https://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/>

Heel snel om te trainen

Haalt goede scores

Decision Trees → heel snel te trainen

When should I use XGBoost?

EXTREME GRADIENT BOOSTING WITH XGBOOST



Sergey Fogelson
VP of Analytics, Viacom



<https://www.youtube.com/watch?v=VgDe0gwesrw>

Features of XGBoost

- Regularized boosting (prevents overfitting)
- Can handle missing values automatically
- Parallel processing
- Can cross-validate at each iteration
 - Enables early stopping, finding optimal number of iterations
- Incremental training
- Can plug in your own optimization objectives
- Tree pruning
 - Generally results in deeper, but optimized, trees

