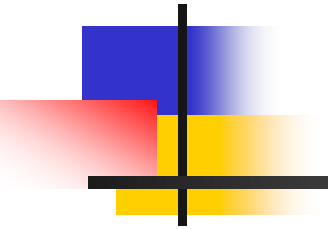
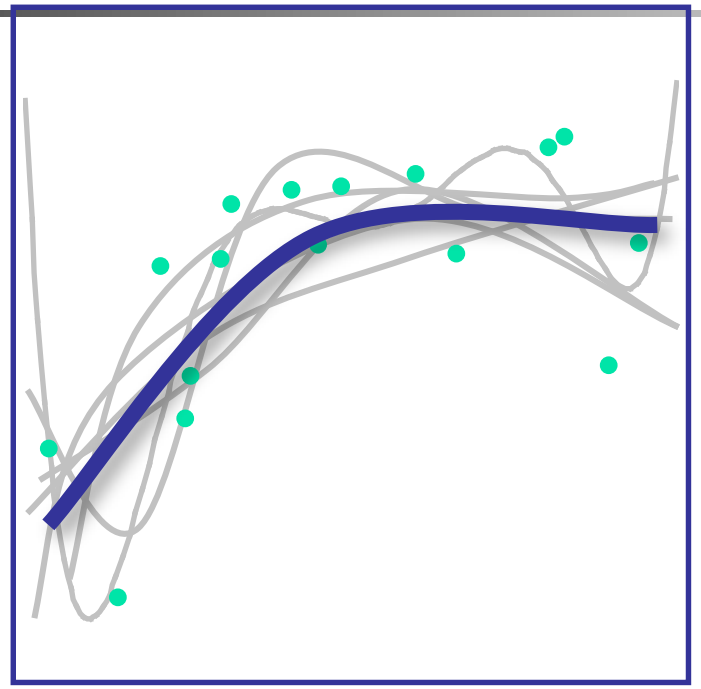


Ensemble Model

Dr. Goutam Chakraborty

Ensemble Models



Combine predictions from multiple models to create a single consensus prediction.

...

How to Combine Models?

- For *prediction estimates (interval targets)* often it's averaging of predictions by all models or, maximum across all models
  - In SAS EM, default for interval targets is Average of Predicted Values
- For *decision (class targets)*, it is often done by voting
  - The *average method in voting* averages the prediction estimates (posterior probabilities) from the models that decide the primary outcome and ignores any model that decides the secondary outcome.
  - The *proportion method in voting* ignores the prediction estimates and instead returns the proportion of models deciding the primary outcome.
  - In SAS EM, default for class targets is Average of Posterior Probabilities.
    - You can change to maximum or Voting
    - If you change to voting, then you can use average or proportion for voting posterior probabilities

Demo

- Delete all connections to the Model Comparison node.
- Click the **Utility** tab. Drag a **Control Point** tool into the diagram workspace and place it to the left of the Model Comparison node.
- Connect model nodes of interest (Misclassification Tree, Probability Tree, Regression (optimal) and Neural Network(Var Selection) to the **Control Point** node. Connect the **Control Point** node to the **Model Comparison** node.
  - The Control Point node serves as a junction in the diagram. A single connection out of the Control Point node is equivalent to all the connections into the node.
- Click the **Model** tab.
- Drag an **Ensemble** tool into the diagram.
- Connect the **Control Point** node to the **Ensemble** node.
- Connect the **Ensemble** node to the **Model Comparison** node
- Run the Model Comparison node and view the results.

More on the Principle of Ensemble

- Ensemble works the best when models going into the Ensemble are very different in nature and produce different predictions.
- In SAS EM you are restricted to combining using specific ways as allowed by the interface.
- But, what if I want to be creative?
  - Such as devise a weighting scheme to give higher weight to models making correct predictions for an observation and lesser weight to models making incorrect prediction for that observation
  - How would I do that?