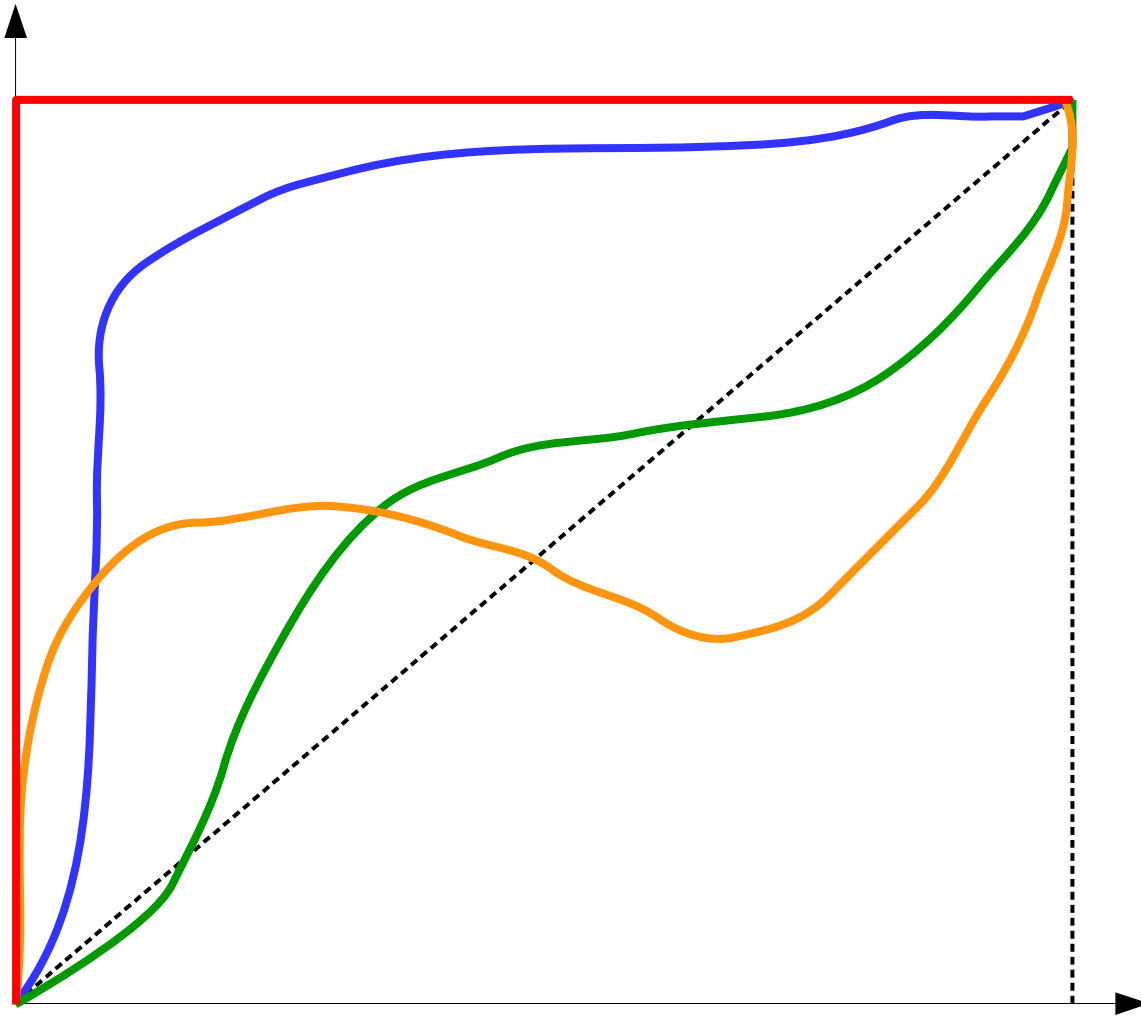


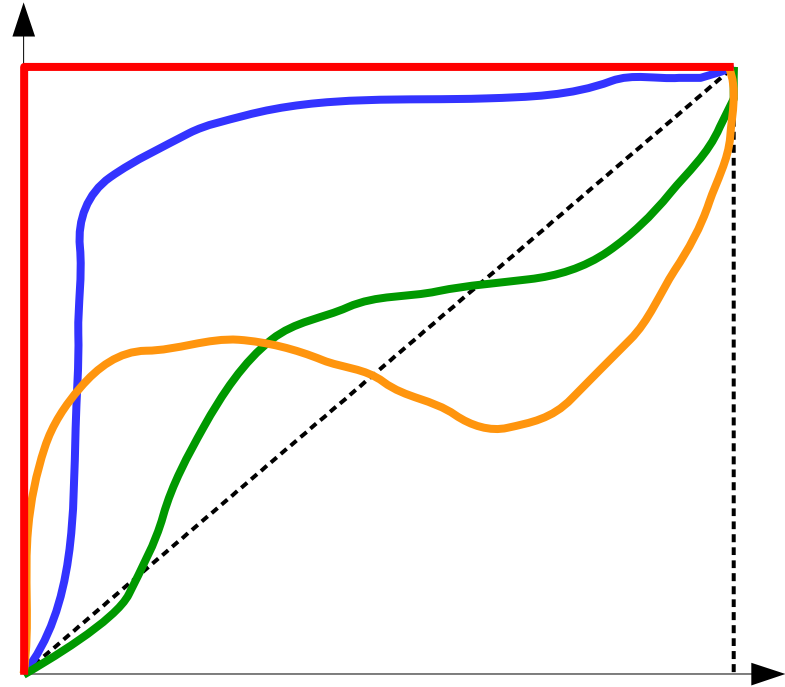
# Curves

- Which of the following curves could be a ROC?
- Which could be a Lift chart?



# Curves

- Which of the following curves could be a ROC?
  - Answer: all, excepted the orange one: TPR and FPR (on the axes) never decrease
- Which could be a Lift chart?
  - Answer: as for ROC, but now also the red one is impossible: you need to classify as positive several records (X axis) to reach 100% of TPR (Y axis)



# Bagging

- We have 3 independent models for the same data, with poor performances
  - Error1 = 45%
  - Error2 = 40%
  - Error3 = 35%
- Is it better to use Model3 alone or to make bagging with all the three models?

# Bagging







- TODO: compute the probability of error of the ensemble
- Standard formula (case for 25 models):

$$\sum_{i=13}^{25} \binom{25}{i} \varepsilon^i (1 - \varepsilon)^{25-i}$$

- Implicitly enumerates all cases with more errors than correct answers ( $i \geq 13$  errors against  $25-i \leq 12$  correct ones)
  - However, it works only when all models have the same error  $\varepsilon$
- Here we have to explicitly enumerate all cases







# Bagging

- Probability of success  and failure  of each:

– Model1 =		.45		.55
– Model2 =		.40		.60
– Model3 =		.35		.65

# Bagging

- We have 8 possible cases

- Model1 =  .45       .55  
- Model2 =  .40       .60  
- Model3 =  .35       .65

Model1



Model2









Model3



# Bagging

- We have 8 possible cases

- Model1 =  .45       .55  
- Model2 =  .40       .60  
- Model3 =  .35       .65

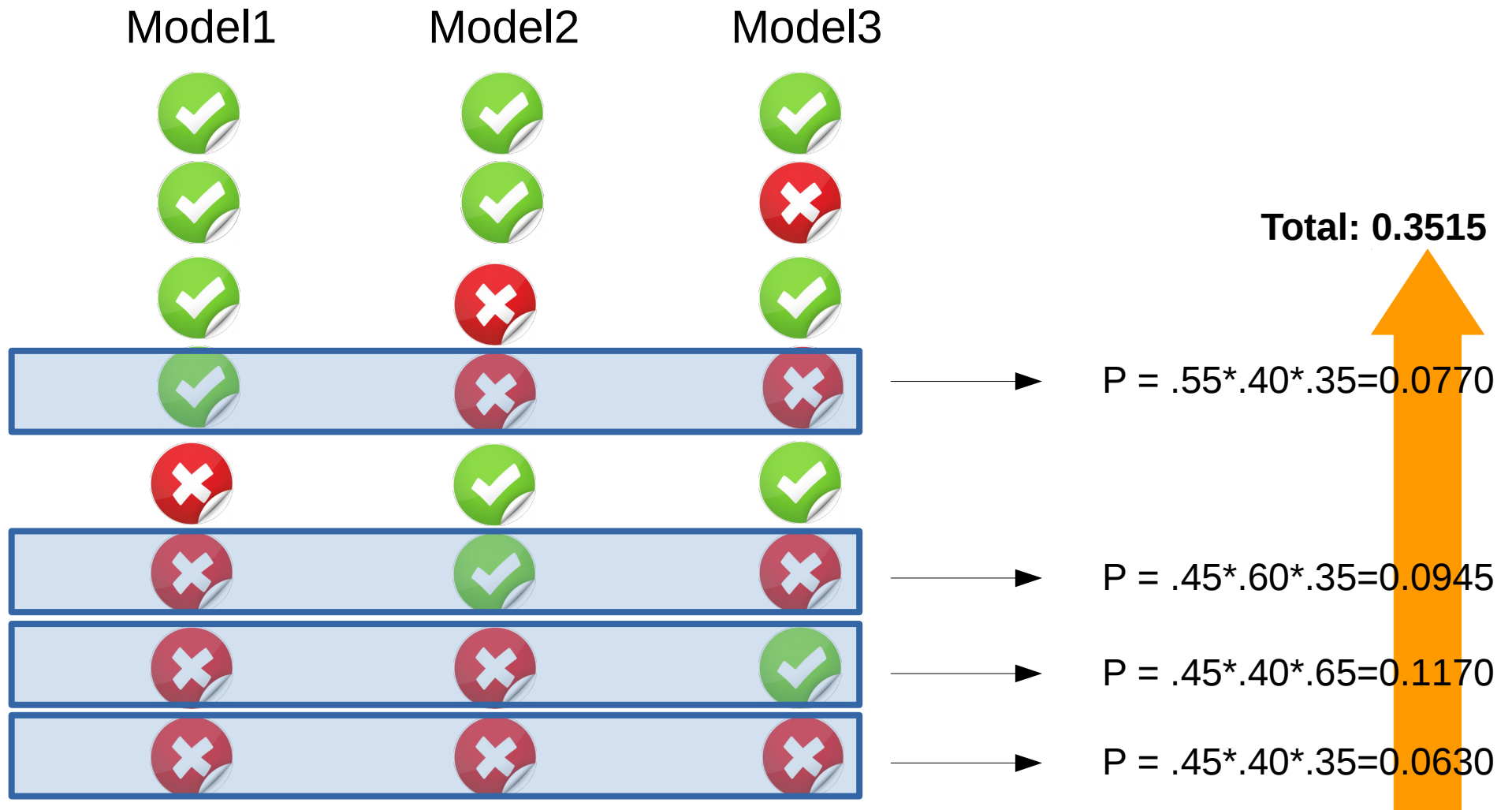


In these cases  
the Bagging  
makes errors

# Bagging

- We have 8 possible cases

- Model1 = .45      .55  
- Model2 = .40      .60  
- Model3 = .35      .65





# Bagging

- Outcome:
  - The “expert” model (Model3) has 35% of error
  - The bagging model has 35.15% of error
  - In this specific case Bagging is not better than the “expert” alone...

# Wisdom of the crowd

- Which of the following question-answering methods might work, thanks to the wisdom of the crowd? Why?
  - Predict the precipitation (rain) level in Sydney by asking to all my neighbors
  - Estimate the average length of taking a degree in CS by asking to all the people in this room
  - Writing an high quality review of a movie by asking to several friends who saw only the trailer
  - Understand if vaccination is good by asking to all my contacts on Facebook (several hundreds)

# Keep an eye on requirements...

- **Diversity of opinion.**
  - People in crowd should have a range of experiences, education and opinions. ( Encourages independent predictions)
- **Independence.**
  - Prediction by person in crowd is not influenced by other people in the crowd.
- **Decentralization.**
  - People have specializations and local knowledge.
- **Aggregation.**
  - There is a mechanism for aggregating all predictions into one single prediction.