**What is Bayes’ Theorem? How is it useful in a machine learning context?**

Bayes’ Theorem gives you the posterior probability of an event given what is known as prior knowledge.

Mathematically, it’s expressed as the true positive rate of a condition sample divided by the sum of the false positive rate of the population and the true positive rate of a condition. Say you had a 60% chance of actually having the flu after a flu test, but out of people who had the flu, the test will be false 50% of the time, and the overall population only has a 5% chance of having the flu. Would you actually have a 60% chance of having the flu after having a positive test?

Bayes’ Theorem says no. It says that you have a (.6 \* 0.05) (True Positive Rate of a Condition Sample) / (.6\*0.05)(True Positive Rate of a Condition Sample) + (.5\*0.95) (False Positive Rate of a Population)  = 0.0594 or 5.94% chance of getting a flu.

**How would you implement a recommendation system for our company’s users?**

Store a set of keywords alongside each product, which should essentially be everything in the title besides a set of stop words When a title is displayed, you find any other products which share keywords in common (with those with one or more in common given priority).

You could further enhance this by assigning a score to each keyword based on its scarcity (with more scarce words being given a higher score, as a match on 'PHP', for instance, is going to be more relevant than a match on 'programming'), or by tracking the number of times a user navigates manually between a set of products.

Regardless you'd best start off by making it simple, and then enhance it as you go on. Depending on the size of your database more advanced techniques may not be all that fruitful.

**Name an example where ensemble techniques might be useful.**

Ensemble techniques use a combination of learning algorithms to optimize better predictive performance. They typically reduce overfitting in models and make the model more robust (unlikely to be influenced by small changes in the training data).

You could list some examples of ensemble methods, from bagging to boosting to a “bucket of models” method and demonstrate how they could increase predictive power.

What Is Deep Learning?

[Deep learning](https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/#48036488d4ba) is a subset of machine learning that involves systems that think and learn like humans using artificial neural networks. The term ‘deep’ comes from the fact that you can have several layers of neural networks.

One of the primary differences between machine learning and deep learning is that feature engineering is done manually in machine learning. In the case of deep learning, the model consisting of neural networks will automatically determine which features to use (and which not to use).

### What Is a Random Forest?

A ‘random forest’ is a supervised machine learning algorithm that is generally used for classification problems. It operates by constructing multiple decision trees during the training phase. The random forest chooses the decision of the majority of the trees as the final decision.

### What Is Decision Tree Classification?

A decision tree builds classification (or regression) models as a tree structure, with datasets broken up into ever-smaller subsets while developing the decision tree, literally in a tree-like way with branches and nodes. Decision trees can handle both categorical and numerical data.