* Gaussian Naive Bayes (GaussianNB)
  + Describe one real-world application in industry where the model can be applied.
    - Naïve Bayes is best known for use in spam filtering, using words in emails to discern the likelihood that an email in question is spam.
  + What are the strengths of the model; when does it perform well?
    - Naïve Bayes computes quickly even when using a lot of variables, and automatically updates with new data, without the model having to be rebuilt. It’s computationally easy to understand and easy to write.
  + What are the weaknesses of the model; when does it perform poorly?
    - Naïve Bayes assumes all variables are independent of each other, which is not always the case.
  + What makes this model a good candidate for the problem, given what you know about the data?
    - Naïve Bayes can handle a large number of variables well while still saving the school a lot on computational costs. It is also cheap to update and modify as new information pours in.
* Decision Trees - http://scikit-learn.org/stable/modules/tree.html
  + Describe one real-world application in industry where the model can be applied.
    - Decision trees are typically used in data mining, for example, if a bank has a giant list of potential loan customers, and must decide if they are credit worthy.
  + What are the strengths of the model; when does it perform well?
    - It’s extremely easy to interpret through which steps a student is classified, which could be important for targeting solutions to factors correlated to dropout risk, and for legal reasons to show students flagged as “at risk” aren’t being unfairly being discriminated against by the algorithm. They also do pretty well with some missing data, with outliers, and with nonlinear relationships.
  + What are the weaknesses of the model; when does it perform poorly?
    - Trees are easy to overfit, and can get quite complex. It’s also better for categorical and discrete variables rather than anything more continuous, like number of absences. New data could easily alter the underlying tree, and since the tree makes decisions on the locally optimal split at each step in the process, it’s unclear that the resulting tree is actually the best possible tree.
  + What makes this model a good candidate for the problem, given what you know about the data?
    - Most of the data seems categorical, and some continuous data such as age are within a small enough range they can be treated categorically. It requires the school to do minimal data cleaning or preparation before executing the algorithm.
* Ensemble Methods (Bagging, AdaBoost, Random Forest, Gradient Boosting) - http://scikit-learn.org/stable/modules/ensemble.html
  + Describe one real-world application in industry where the model can be applied.
    - Ensemble methods are often used in facial recognition technology.
  + What are the strengths of the model; when does it perform well?
    - Ensemble methods often have high accuracy, and focus on the data they fail to categorize correctly, working on their “weak spots.”
  + What are the weaknesses of the model; when does it perform poorly?
    - Ensemble methods are sensitive to outliers and noise in the data, trying to overcompensate to fit the noise.
  + What makes this model a good candidate for the problem, given what you know about the data?
    - This model is highly accurate, so meets the school’s specifications there, though it can be computationally intensive.
* K-Nearest Neighbors (KNeighbors)
  + Describe one real-world application in industry where the model can be applied.
    - A version of KNN, I believe, is used by Netflix to recommend movies and TV shows to users based on their past choices and the choices of people with tastes similar to them.
  + What are the strengths of the model; when does it perform well?
    - KNN is typically pretty accurate. It also has extremely low up-front data costs, though each subsequent computation will require some cost.
  + What are the weaknesses of the model; when does it perform poorly?
    - The mechanics behind a KNN classification are typically hard to interpret. It’s also prone to overfitting, with models reaching a point at which increasing accuracy on the training set is inverse to how well it generalizes to outside data. KNN also operates off of the concept of “distance,” meaning it does better with continuous variables than anything categorical.
  + What makes this model a good candidate for the problem, given what you know about the data?
    - This model will find students whose data most closely resembles students who have dropped out, finding interactions between variables that might be hard for other models to pick up on.
* Stochastic Gradient Descent (SGDC)
  + Describe one real-world application in industry where the model can be applied.
  + What are the strengths of the model; when does it perform well?
  + What are the weaknesses of the model; when does it perform poorly?
  + What makes this model a good candidate for the problem, given what you know about the data?
* Support Vector Machines (SVM)
  + Describe one real-world application in industry where the model can be applied.
  + What are the strengths of the model; when does it perform well?
    - SVMs are generally quite accurate, comparable to KNN models. They can also separate data non-linearly. Finally, SVMs operate well in high dimensions.
  + What are the weaknesses of the model; when does it perform poorly?
    - SVM does better with continuous variables than discrete variables, and it very computationally costly.
  + What makes this model a good candidate for the problem, given what you know about the data?
    - SVMs may pick up on non-linear separation in the data, and have high accuracy.
* Logistic Regression
  + Describe one real-world application in industry where the model can be applied.
    - Logistic regression can be used in product testing, to establish thresholds for same use of the product in question.
  + What are the strengths of the model; when does it perform well?
    - Logistic regression is extremely efficient and low-cost when categorizing a new student, once the underlying algorithm has been decided. The output can also easily be interpreted as a probability, rather than a yes/no, which can help the school decide on which students are most in danger of failing.
  + What are the weaknesses of the model; when does it perform poorly?
    - Logistic regression makes some of the same assumptions about the data as linear regression (i.e. error terms need to be independent, you can easily overfit by adding in too many variables, etc.) and typically is not as accurate as many other methods.
  + What makes this model a good candidate for the problem, given what you know about the data?
    - Logistic regression is simple, cheap, and provides a good baseline off of which to explore other models.