List three of the supervised learning models above that are appropriate for this problem that you will test on the census data. For each model chosen:

* *Describe one real-world application in industry where the model can be applied.* (You may need to do research for this — give references!)
* *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?*
* **Decision Trees:** 
  + Example: Medical diagnoses: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.109.7119&rep=rep1&type=pdf>
    - http://www.cbcb.umd.edu/~salzberg/docs/murthy\_thesis/survey/node32.html
  + Strengths:
    - Works well for general classification problems
    - It is robust to errors and data omissions
    - Does best with:
      * discrete variables
      * Instances are represented by feature-value pairs (i.e. Temperature 🡪 Hot, Cold)
      * Disjunctive data
    - http://www.cs.princeton.edu/courses/archive/spr07/cos424/papers/mitchell-dectrees.pdf
  + Weaknesses:
    - Continuous data
    - Interrelated or connected data
  + Because the data in the CharityML problem is discrete, and the goal is to make a decision on whether to contact someone for a donation, this fits the bill of a classification problem that fits well with the strengths listed above for Decision Trees.
* **Ensemble Methods (Bagging, AdaBoost, Random Forest, Gradient Boosting)**
  + Example: Use in image classification, such as identifying street signs: <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx2aGRiYWxhbGl8Z3g6NjU1MTlmYzI0MTkwY2E4Yg>
  + Strengths:
    - versatile – can be made up of a mix of homogeneous or heterogeneous learners
    - Ensemble methods are good classification tools. They are good at whatever the individual learners that make up the ensemble are good at
    - https://www.islab.ntua.gr/attachments/article/86/Ensemble%20methods%20-%20Zhou.pdf
  + Weakneses:
    - Computation cost is higher than a single learner (only marginally higher than training an individual learner)
  + Ensemble methods are a good candidate for this problem due to the categorical nature of the problem.
* **Support Vector Machines (SVM)**
  + SVM can be used in image recognition algorithms, such as in identifying street signs
    - <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx2aGRiYWxhbGl8Z3g6NjU1MTlmYzI0MTkwY2E4Yg>
  + Strengths
    - It works well when there is a clear margin of separation
    - It is memory efficient
    - It works well with high dimensions
  + Weaknesses
    - It is not good for large datasets because of the slow processing time
    - It does not perform well with lots of noise
  + <https://www.analyticsvidhya.com/blog/2015/10/understaing-support-vector-machine-example-code/>
  + SVM’s are a good candidate due to the small data size and multiple dimensions to the data
* K-Nearest Neighbors (KNeighbors)
* Stochastic Gradient Descent Classifier (SGDC)
  + Good for extremely large data sets – highly efficient
* Logistic Regression
  + Best with conditional probabilities
  + Not with yes/no answers but with “how yes” and “how no”
* GuassianNB:
  + Best for: text classification