1. Drop rows with any missing value:

```
print len(loans[[target] + features])
print len(loans[[target] + features].dropna(axis=0, how="any"))
```

```
122607
122578
```

2. Return rows with any missing value:

```
df = loans[[target] + features]
len(df[df.isnull().any(axis=1)])
```

```
29
```

3. Ada Boost Classifier:

```
from sklearn.ensemble import AdaBoostClassifier
```

```
clf = AdaBoostClassifier(n_estimators = 5)
model_5 = clf.fit(train_data.drop(target, axis=1), train_data[target])
```

4. Predict class:

```
model_5.predict(sample_validation_data.drop(["safe_loans"], axis=1))
```

```
array([-1,  1, -1, -1], dtype=int64)
```

5. Predict probabilities:

```
model_5.predict_proba(sample_validation_data.drop(["safe_loans"], axis=1))[:, 1]
```

```
array([0.49569827, 0.54590972, 0.47792678, 0.49569827])
```