Notes:

1 hour? 40 minutes?

Idiots -> Experts?

How much column generation?

Both formulations of RMP?

Both methods of calculating rank?

# Motivation

## What is classification

## Why do we want better classifiers

## Things that can confuse a classifier

# Instance selection

## Explanation

## Previous work

## Instance selection as an integer program

### Column generation about

### Column generation specifics

### Overall algorithm

# Results

## Experimental

### Overall

### Overfitting

### What it is doing when it works well

#### overlap

#### outliers

#### minority class

### Comparison

#### hard to pick

#### simple and easy

## Case Study

#### Study 1

#### Study 2

#### Rules of thumb?

# Conclude

## Impact

## Future work