Predicting offending on bail

Felix Leung, BOCSAR
Outline
Outline

1. Aim: to explore the potential use of predictive modelling to better inform bail decisions
Outline

1. Aim: to explore the potential use of predictive modelling to better inform bail decisions
2. Predictive modelling
Outline

1. Aim: to explore the potential use of predictive modelling to better inform bail decisions
2. Predictive modelling
3. Results
1. Aim
Bail outcomes
Bail outcomes

1. Bail dispensed with
Bail outcomes

1. Bail dispensed with

2. Bail granted
Bail outcomes

1. Bail dispensed with

2. Bail granted

3. Bail refused
Bail outcomes

1. Bail dispensed with

2. Bail granted

3. Bail refused
Bail outcomes in 2016

- Bail dispensed with: 89,703
- Bail granted: 30,334
- Bail refused: 12,143

Year: 2016

Range: 0 to 100,000
Bail outcomes in 2016

Year 2016

Bail granted - Did not offend: 21330
Bail granted - Offended: 9004
Bail refused: 12143

Legend:
- Blue: Bail granted - Did not offend
- Orange: Bail granted - Offended
- Black: Bail refused
Aim
Aim

Specifically, could we:
Aim

Specifically, could we:

1. Reduce **offending on bail** (without increasing **bail refusals**)?
Aim

Specifically, could we:

1. Reduce **offending on bail** (without increasing **bail refusals**)?
2. Reduce **bail refusals** (without increasing **offending on bail**)?
Aim

Specifically, could we:

1. Reduce **offending on bail** (without increasing **bail refusals**)?
2. Reduce **bail refusals** (without increasing **offending on bail**)?

![Graph showing bail outcomes for 2016]

- **Bail granted - Did not offend**: 21,330
- **Bail granted - Offended**: 9,004
- **Bail refused**: 12,143

Year 2016
2. Predictive modelling
Data

- NSW Reoffending Database
Data

- NSW Reoffending Database
- Include:
  - Defendants charged between 2011 and 2016
Data

- NSW Reoffending Database
- Include:
  - Defendants charged between 2011 and 2016
- Exclude:
  - ANZSOC subdivision 152 Breach of community-based order
  - ANZSOC group 1569 Offences against justice procedures, nec

(ANZSOC: Australian and New Zealand Standard Offence Classification)
Some terminology
Some terminology

- **Outcome**: new proven offence committed on bail
Some terminology

- **Outcome**: new proven offence committed on bail
- **Predictors**: variables available before charge date
Some terminology

- **Outcome**: new proven offence committed on bail
- **Predictors**: variables available before charge date
- **Model**: approximate relationship between outcome and predictors
Some terminology

- **Outcome**: new proven offence committed on bail
- **Predictors**: variables available before charge date
- **Model**: approximate relationship between outcome and predictors
- **Training data**: between 2011 and 2015
Some terminology

- **Outcome**: new proven offence committed on bail
- **Predictors**: variables available before charge date
- **Model**: approximate relationship between outcome and predictors
- **Training data**: between 2011 and 2015
- **Test data**: 2016
Predictors

1. Demographics:
   - e.g. age, gender, socio-economic index (SEIFA)
Predictors

1. Demographics:
   - e.g. age, gender, socio-economic index (SEIFA)

2. Variables on index offence:
   - e.g. whether offence indictable, summary, DV-related
   - ANZSOC divisions and subdivisions
Predictors

1. Demographics:
   - e.g. age, gender, socio-economic index (SEIFA)

2. Variables on index offence:
   - e.g. whether offence indictable, summary, DV-related
   - ANZSOC divisions and subdivisions

3. Criminal history:
   - Age at first criminal justice contact, no. guilty charges
   - Previous penalty: prison, fine, bond, etc
   - and more
3 Models compared
3 Models compared

(MARS: multivariate adaptive regression splines (Friedman, 1991))
3 Models compared

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Linear (logistic)</th>
<th>Non-linear (MARS, d=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictor 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictor 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Bail granted - Did not offend
- Bail granted - Offended

(MARS: multivariate adaptive regression splines (Friedman, 1991))
3 Models compared

Linear (logistic) | Non-linear (MARS, d=1) | Non-additive (MARS, d=2)

- Bail granted - Did not offend
- Bail granted - Offended

(MARS: multivariate adaptive regression splines (Friedman, 1991))
Model selection

based on AUC (area under curve)
Model selection

based on AUC (area under curve)

AUC

Linear

Non-linear

Non-additive
Predictors selected
Predictors selected

1. Demographics:
   - Age at charge date
Predictors selected

1. Demographics:
   - Age at charge date

2. Variables on index offence:
   - Median Severity Ranking of most serious offence charged
   - Most serious offence in ANZSOC 03 (Sexual assault)
   - Most serious offence in ANZSOC 08 (Theft)
   - Most serious offence in ANZSOC 151 (Breach custodial order)
Predictors selected

1. Demographics:
   - Age at charge date

2. Variables on index offence:
   - Median Severity Ranking of most serious offence charged
   - Most serious offence in ANZSOC 03 (Sexual assault)
   - Most serious offence in ANZSOC 08 (Theft)
   - Most serious offence in ANZSOC 151 (Breach custodial order)

3. Criminal history:
   - Age at first criminal justice contact
   - No. of prev. charges
   - No. of prev. prison
   - Proportion of prev. bail with new proven offence
   - Time since prev. charge
3. Results
Some remarks
Some remarks

1. The model predictions are in terms of probabilities, i.e. likelihood of offending
Some remarks

1. The model predictions are in terms of probabilities, i.e. likelihood of offending

2. To compare model predictions with 2016 figures, some threshold is needed
Some remarks

1. The model predictions are in terms of probabilities, i.e. likelihood of offending

2. To compare model predictions with 2016 figures, some threshold is needed

3. Here we assume hypothetically:
   - individual **above** the threshold would be refused bail
   - individual **below** the threshold would be granted bail
Some remarks

1. The model predictions are in terms of probabilities, i.e. likelihood of offending.

2. To compare model predictions with 2016 figures, some threshold is needed.

3. Here we assume hypothetically:
   - individual **above** the threshold would be refused bail
   - individual **below** the threshold would be granted bail

Bail outcomes in 2016

Year: 2016

- Bail granted - Did not offend: 21330
- Bail granted - Offended: 9004
- Bail refused: 12143

Legend:
- Blue: Bail granted - Did not offend
- Orange: Bail granted - Offended
- Black: Bail refused
Model Predictions

- **21330**: Bail granted - Did not offend
- **9004**: Bail granted - Offended
- **12143**: Bail refused
Model Predictions

- **21330**
  - Bail granted - Did not offend
- **9004**
  - Bail granted - Offended
- **12143**
  - Bail refused
Model Predictions

Bail granted - Did not offend: 21330
Bail granted - Offended: 9004
Bail refused: 12143

Histograms showing the likelihood of offending for each category.
Aim

Could we:

1. Reduce **offending on bail** (without increasing **bail refusals**)?
2. Reduce **bail refusals** (without increasing **offending on bail**)?
1. Reduce offending on bail

- 21330: Bail granted - Did not offend
- 9004: Bail granted - Offended
- 12143: Bail refused
1. Reduce **offending on bail**

- **21330**
  - Bail granted - Did not offend

- **9004**
  - Bail granted - Offended

- **12143**
  - Bail refused

---

**Likelihood of offending**

- Count

**Likelihood of offending**

- Count
1. Reduce offending on bail

- **21330**
  - Bail granted - Did not offend

- **9004**
  - Bail granted - Offended

- **12143**
  - Bail refused
1. Reduce **offending on bail**

- **Bail granted - Did not offend**
  - Count: 21330
  - Likelyhood of offending:
    - Lower than 0.25: 1000
    - Between 0.25 and 0.50: 1000
    - Between 0.50 and 0.75: 1000
    - Greater than 0.75: 2000

- **Bail granted - Offended**
  - Count: 9004
  - Likelyhood of offending:
    - Lower than 0.25: 1000
    - Between 0.25 and 0.50: 2000
    - Between 0.50 and 0.75: 0
    - Greater than 0.75: 2000

- **Bail refused**
  - Count: 12143
  - Likelyhood of offending:
    - All likelihood ranges have the same count of 1000

**Hypothetically bail granted - offended:**

- Total count: 8304
- Threshold: 30/45
1. Reduce **offending on bail**

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
<th>Likelihood of Offending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bail granted</td>
<td>21330</td>
<td></td>
</tr>
<tr>
<td>Did not offend</td>
<td></td>
<td>0.00 - 0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Refused</td>
</tr>
<tr>
<td>Bail granted</td>
<td>9004</td>
<td></td>
</tr>
<tr>
<td>Offended</td>
<td></td>
<td>0.25 - 0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Refused</td>
</tr>
<tr>
<td>Bail refused</td>
<td>12143</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hypothetically refused</td>
</tr>
</tbody>
</table>

Hypothetically bail refused: 12140
1. Reduce offending on bail

- **21330**
  - Bail granted - Did not offend

- **9004**
  - Bail granted - Offended

- **12143**
  - Bail refused

Hypothetically bail refused: 12140

Below: 19377
Offended: 0
1. Reduce **offending on bail**

- **21330**
  - Bail granted - Did not offend

- **9004**
  - Bail granted - Offended

- **12143**
  - Bail refused

Below: 19377
Offended: 0

Hypothetically bail refused: 12140

Below: 5648
Offended: 5648
1. Reduce **offending on bail**

<table>
<thead>
<tr>
<th>Bail granted - Did not offend</th>
<th>Bail granted - Offended</th>
<th>Bail refused</th>
</tr>
</thead>
<tbody>
<tr>
<td>21330</td>
<td>9004</td>
<td>12143</td>
</tr>
</tbody>
</table>

**Hypothetically bail refused**: 12140

- Below: 19377
- Offended: 0

- Below: 5648
- Offended: 5648

- Below: 5312
- Offended: ?
1. Reduce **offending on bail**

**Bail granted - Did not offend**
- Count: 21330

**Bail granted - Offended**
- Count: 9004

**Bail refused**
- Count: 12143

**Hypothetically bail refused: 12140**
- Below: 19377
  - Offended: 0
- Below: 5648
  - Offended: 5648
- Below: 5312
  - Offended: 2656

(assumed 50% would have)
1. Reduce offending on bail

**Bail granted - Did not offend**: 21330

**Bail granted - Offended**: 9004

**Bail refused**: 12143

Hypothetically bail refused: 12140
Hypothetically bail granted - offended: 5648 + 2656 = 8304
1. Reduce offending on bail

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Likelihood of Offending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bail granted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did not offend</td>
<td>21330</td>
<td></td>
</tr>
<tr>
<td>Offended</td>
<td>9004</td>
<td></td>
</tr>
<tr>
<td>Bail refused</td>
<td>12143</td>
<td></td>
</tr>
</tbody>
</table>

Hypothetically bail refused: 12140
Hypothetically bail granted - offended: 5648 + 2656 = 8304
(a reduction of 700 or 7.8%)
2. Reduce bail refusals

- **21330**
  - Bail granted - Did not offend
- **9004**
  - Bail granted - Offended
- **12143**
  - Bail refused

Hypothetically bail granted - offended: 5648 + 2656 = 8304

38 / 45
2. Reduce bail refusals

- **Bail granted - Did not offend**: 21330
- **Bail granted - Offended**: 9004
- **Bail refused**: 12143

Below: 19377
Offended: 0
Below: 6010
Offended: 6010
Below: 5988
Offended: 2994

**Hypothetically bail granted - offended**: 5648 + 2656 = 8304

Below: 19377
Offended: 0
Below: 6010
Offended: 6010
Below: 5988
Offended: 2994

(assumed 50% would have)
2. Reduce bail refusals

21330
Bail granted - Did not offend

9004
Bail granted - Offended

12143
Bail refused

Hypothetically bail granted - offended: 6010 + 2994 = 9004
2. Reduce bail refusals

Bail granted - Did not offend: 21330
Bail granted - Offended: 9004
Bail refused: 12143

Hypothetically bail refused: 10750
Hypothetically bail granted - offened: 6010 + 2994 = 9004
2. Reduce bail refusals

Hypothetically bail refused: 10750

Hypothetically bail granted - offended: $6010 + 2994 = 9004$

(a reduction of 1393 or 11.5%)
Could we:

1. Reduce **offending on bail** (without increasing **bail refusals**)?
2. Reduce **bail refusals** (without increasing **offending on bail**)?
Summary

Could we:

1. Reduce *offending on bail* (without increasing *bail refusals*)?
2. Reduce *bail refusals* (without increasing *offending on bail*)?

What we found:

1. *offending on bail*: 700 or 7.8% reduction
2. *bail refusals*: 1393 or 11.5% reduction

(this is assuming 50% of those *bail refused* but below the threshold would have offended if granted bail.)
Limitations
Limitations

1. Other factors that influence bail decisions are ignored, e.g. likelihood of non-appearance and the severity of the offence
Limitations

1. Other factors that influence bail decisions are ignored, e.g. likelihood of non-appearance and the severity of the offence

2. Offending is measured crudely; some of the offending predicted would not be serious enough to change a bail decision
Limitations

1. Other factors that influence bail decisions are ignored, e.g. likelihood of non-appearance and the severity of the offence

2. Offending is measured crudely; some of the offending predicted would not be serious enough to change a bail decision

3. Other important concerns such as fairness, non-representative data are not addressed
References
