Featurespace

Case study of harm minimisation

2 February 2017

FEATURE SPACE

OUTSMART RISK

Agenda

1. Introduction
2. Data & Labels
3. Modelling
4. Results
5. Production deployment

FEATURE SPACE

OUTSMART RISK

Agenda

1. **Introduction**
2. Data & Labels
3. Modelling
4. Results
5. Production deployment

FEATURE SPACE

OUTSMART RISK

Reduce problem gambling by reducing the stakes on fixed odds machines

Saturday 28 March 2015

HILARY Douglas of the Association of British B

£

£33.2 million: The amount staked on "dangerously addictive" betting machines in Bolton

Betting machines: How one man lost everything, £1,000 at a t

Gambling addict tells of destruction wrought in
odds betting terminals - and battle to change his

GOVERNMENT BRANDED 'BUNCH OF TOSSERS' OVER MACHINE REGULATIONS

Posted by: Andrew McCarron March 30, 2015 in Features, Latest News, Slider Images, UK Comment

Tweet 6 Like 1 +1 1 Share 8

Ahead of a General Election in which none of the parties have garnered

Problem Statement

Is it possible to distinguish between harmful and non-harmful gaming machine play?

responsible gamblingtrust

RGSB Responsible Gambling Strategy Board
RESEARCH ▪ EDUCATION ▪ TREATMENT

GAMBLING COMMISSION

Department for Culture Media & Sport

What is a gaming machine?

Industry Cooperation

Limitations

- **Definition of Harm**: In this project, the PGSI screen has been used as a proxy to identify harm.
- **Defining the unit of analysis as a 'Session'**: The unit of continuous play used in the analysis has been a session. This does not capture a player's entire visit to a venue, which could comprise multiple sessions.
- **Understanding Bet Selection and Gaming Machine Browsing**: Understanding selection of bets on Roulette, or navigation between menus on a gaming machine, would provide further insight.
- **Defining a player and restricted card usage**: Only data associated with a player's card has been analysed. We know some players have multiple cards, and sometimes play without their card.
- **Multiple Gambling Product Engagement**: The players surveyed engage with multiple gambling products. This analysis only looks at their gaming machine play.

Agenda

1. Introduction
2. **Data & Labels**
3. Modelling
4. Results
5. Production deployment

FEATURE SPACE

OUTSMART RISK

Gaming Machine data

PlayerID	Timestamp	Value	Balance	Action	Game
	09:18	-11.60	11.80	Play	Roulette
	09:18	7.20	19.00	Win	Roulette
	09:18	-11.60	7.40	Play	Roulette
	09:18	12.60	20.00	Win	Roulette
	09:19	-11.60	8.40	Play	Roulette
	09:19	14.40	22.80	Win	Roulette
	09:19	-11.60	11.20	Play	Roulette
	09:20	-11.20	0	Play	Roulette
123456	12:53	10.00	10.00	Cash In	
123456	12:53	10.00	20.00	Cash In	
123456	12:54	10.00	30.00	Cash In	
	13:01	-30.00	0	Cash Out	
	13:05	2.00	2.00	Cash In	Roulette
	13:05	0.10	2.10	Cash In	Roulette
	13:05	-2.10	0	Play	Roulette
	13:05	3.60	3.60	Win	Roulette

How much data did we use?

Total number of bets between September 2013 and June 2014

6,768,053,704

We had every single interaction with a gaming machine in the UK for 10 months, over 10 billion of them!

NatCen Social Research

Sessionized data

PlayerID	Timestamp	Value	Balance	Action	Game	Proxy Session Score
	09:18	-11.60	11.80	Play	Roulette	0.00
	09:18	7.20	19.00	Win	Roulette	0.00
	09:18	-11.60	7.40	Play	Roulette	0.00
	09:18	12.60	20.00	Win	Roulette	0.00
	09:19	-11.60	8.40	Play	Roulette	0.00
	09:19	14.40	22.80	Win	Roulette	0.00
	09:19	-11.60	11.20	Play	Roulette	0.00
	09:20	-11.20	0	Play	Roulette	0.00
123456	12:53	10.00	10.00	Cash In		0.58
123456	12:53	10.00	20.00	Cash In		0.08
123456	12:54	10.00	30.00	Cash In		0.04
	13:01	-30.00	0	Cash Out		0.00
	13:05	2.00	2.00	Cash In	Roulette	0.38
	13:05	0.10	2.10	Cash In	Roulette	0.04
	13:05	-2.10	0	Play	Roulette	0.00
	13:05	3.60	3.60	Win	Roulette	0.00

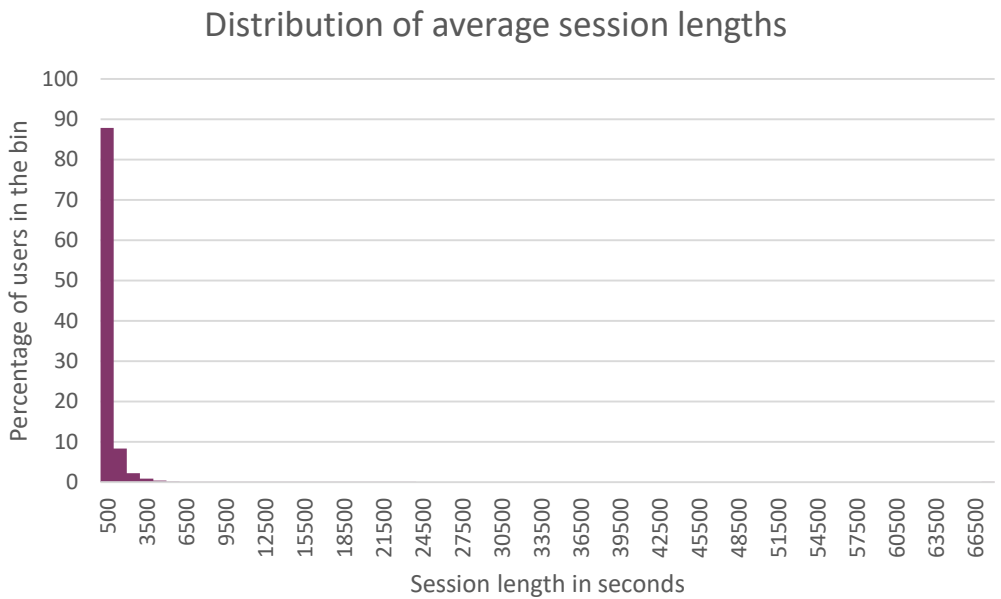
Sessionized Data

PlayerID	Timestamp	Value	Balance	Action	Game	Proxy Session Score	Session ID	Proxy Session PlayerID
	09:18	-11.60	11.80	Play	Roulette	0.00	1	
	09:18	7.20	19.00	Win	Roulette	0.00	1	
	09:18	-11.60	7.40	Play	Roulette	0.00	1	
	09:18	12.60	20.00	Win	Roulette	0.00	1	
	09:19	-11.60	8.40	Play	Roulette	0.00	1	
	09:19	14.40	22.80	Win	Roulette	0.00	1	
	09:19	-11.60	11.20	Play	Roulette	0.00	1	
	09:20	-11.20	0	Play	Roulette	0.00	1	
123456	12:53	10.00	10.00	Cash In		0.58	2	123456
123456	12:53	10.00	20.00	Cash In		0.08	2	123456
123456	12:54	10.00	30.00	Cash In		0.04	2	123456
	13:01	-30.00	0	Cash Out		0.00	2	123456
	13:05	2.00	2.00	Cash In	Roulette	0.38	3	
	13:05	0.10	2.10	Cash In	Roulette	0.04	3	
	13:05	-2.10	0	Play	Roulette	0.00	3	
	13:05	3.60	3.60	Win	Roulette	0.00	3	

Problem gambling severity index

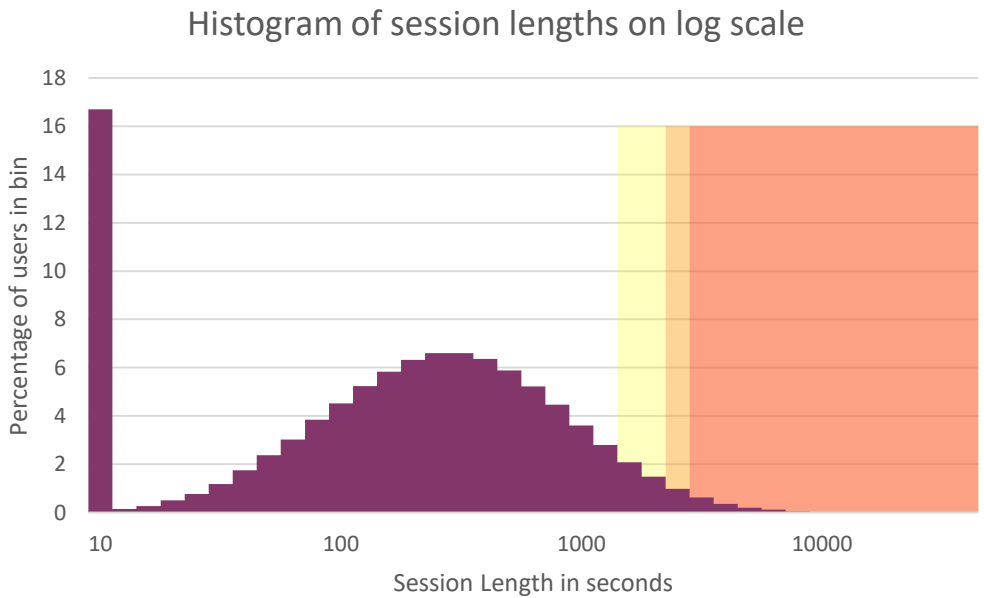
- Comes from Canadian research - 2001
  - When you think of the past 12 months, have you bet more than you could really afford to lose?
  - Still thinking about the last 12 months, have you needed to gamble with larger amounts of money to get the same feeling of excitement?
  - When you gambled, did you go back another day to try to win back the money you lost?
  - Have you borrowed money or sold anything to get money to gamble?
  - Have you felt that you might have a problem with gambling?
  - Has gambling caused you any health problems, including stress or anxiety?
  - Have people criticized your betting or told you that you had a gambling problem, regardless of whether or not you thought it was true?
  - Has your gambling caused any financial problems for you or your household?
  - Have you felt guilty about the way you gamble or what happens when you gamble?
- For each answer:
  - Never = 0 pts
  - Sometimes = 1 pt
  - Most of the time = 2 pts
  - Almost always = 3 pts
- Score of 0: Non-problem gambling.
- Score of 1 or 2: Low level of problems with few or no identified negative consequences.
- Score of 3 to 7: Moderate level of problems leading to some negative consequences.
- Score of 8 or more: Problem gambling with negative consequences and a possible loss of control.

Distributions are very skewed



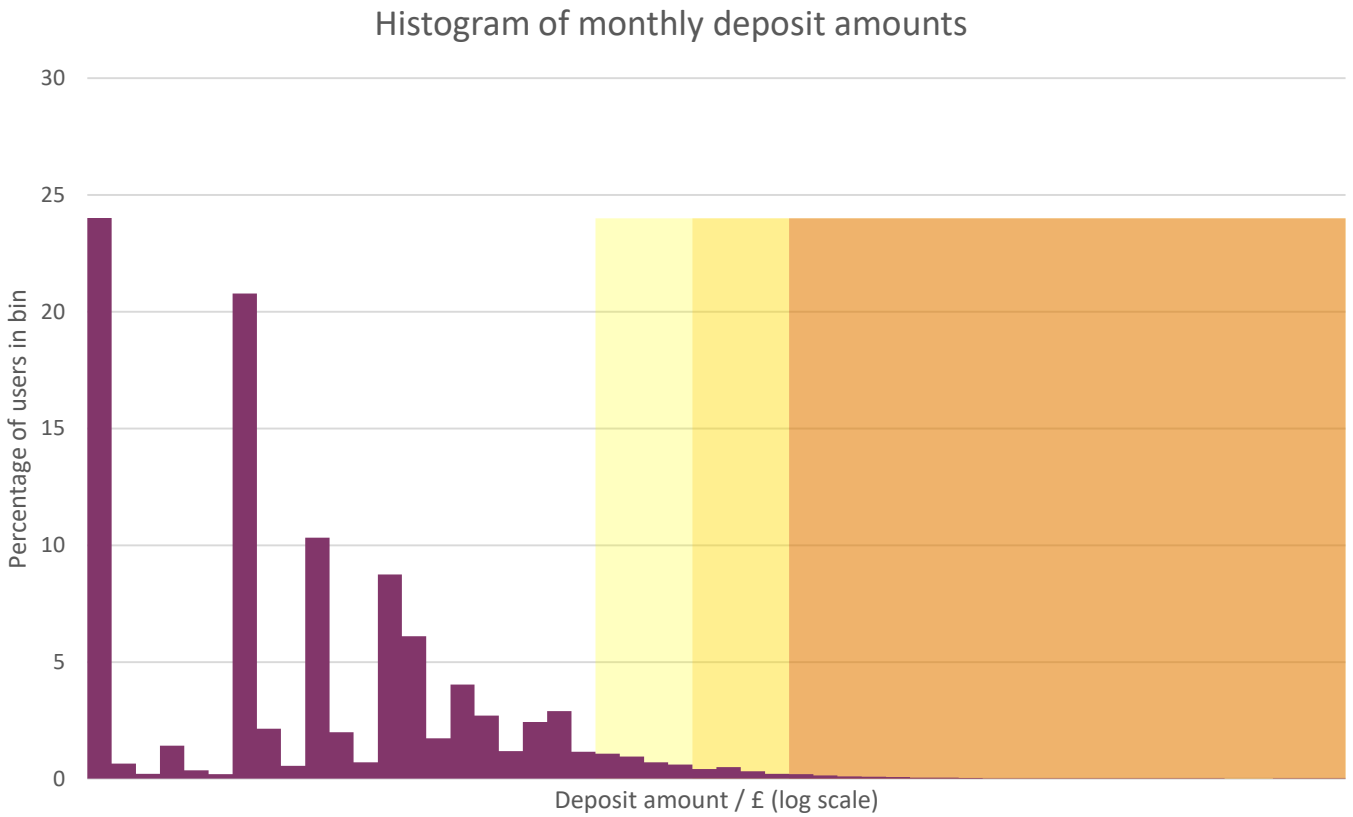
> **The data is very skewed so the distribution is difficult to see**

Log scales help



> **Use a log scale**

Data distribution is definitely driven by people



Agenda

1. Introduction
2. Data & Labels
3. **Modelling**
4. Results
5. Production deployment

FEATURE SPACE

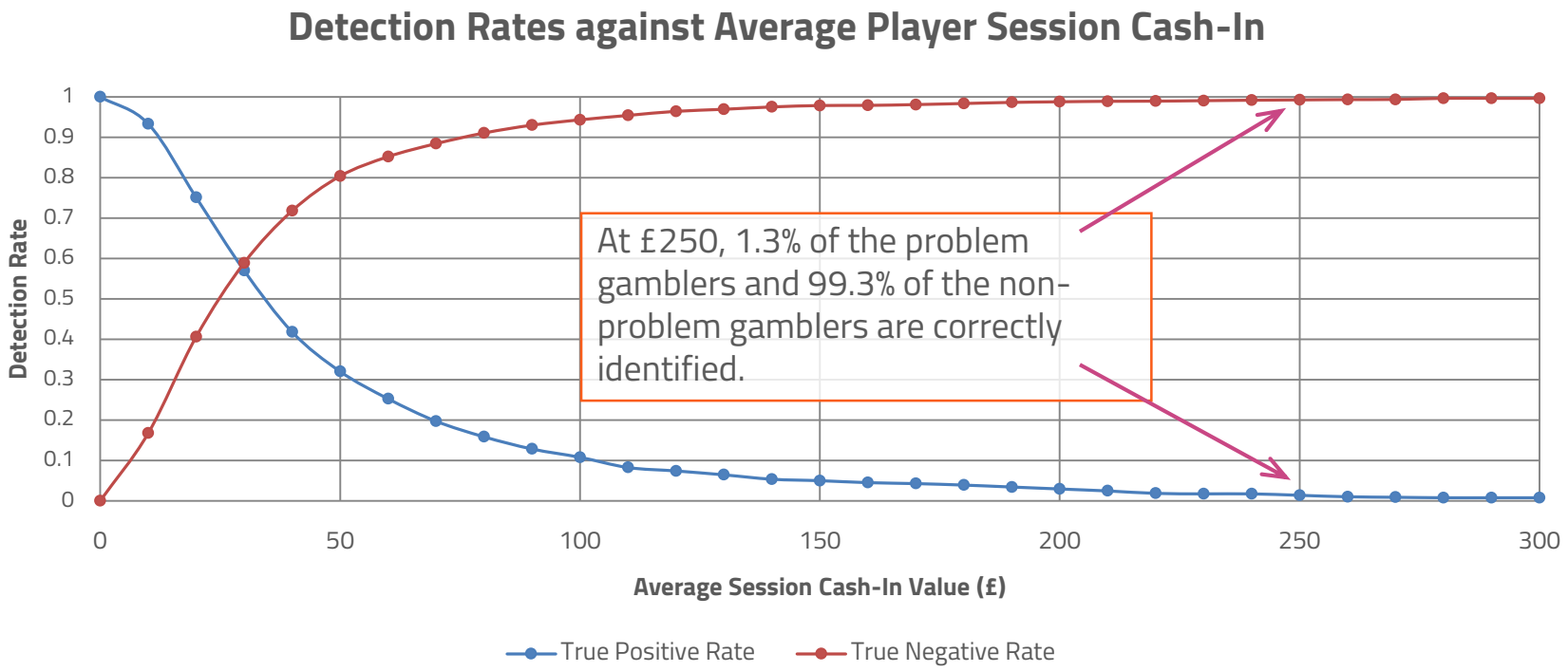
OUTSMART RISK

Modelling

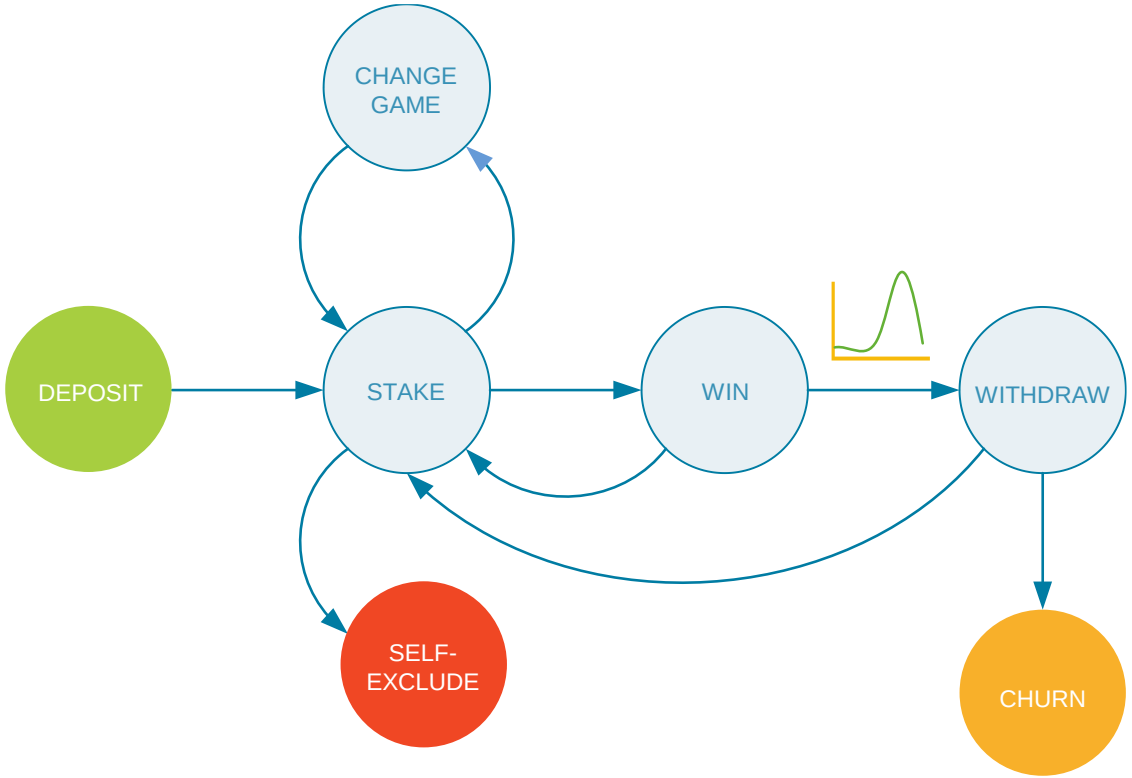
- Baseline model used by industry: Problem gambler if session spend >£250 or session time > 30 min

- We used ‘classical’ approaches to Machine Learning
- Understand the domain
- Enumerate the behaviours that we think are likely to be indicative
- Decide what to compute to capture this behaviour
- Learn a function to generate predictions

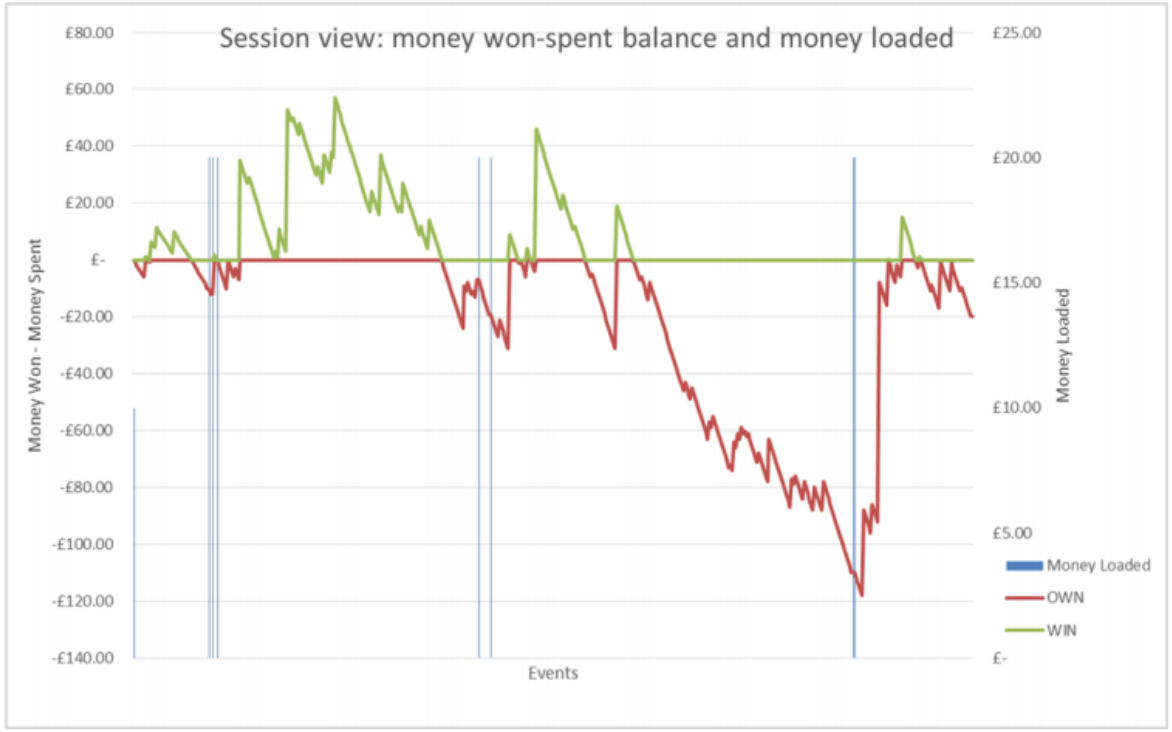
Using player spend



Understanding the player journey



Player behaviour different between staking wins and own money



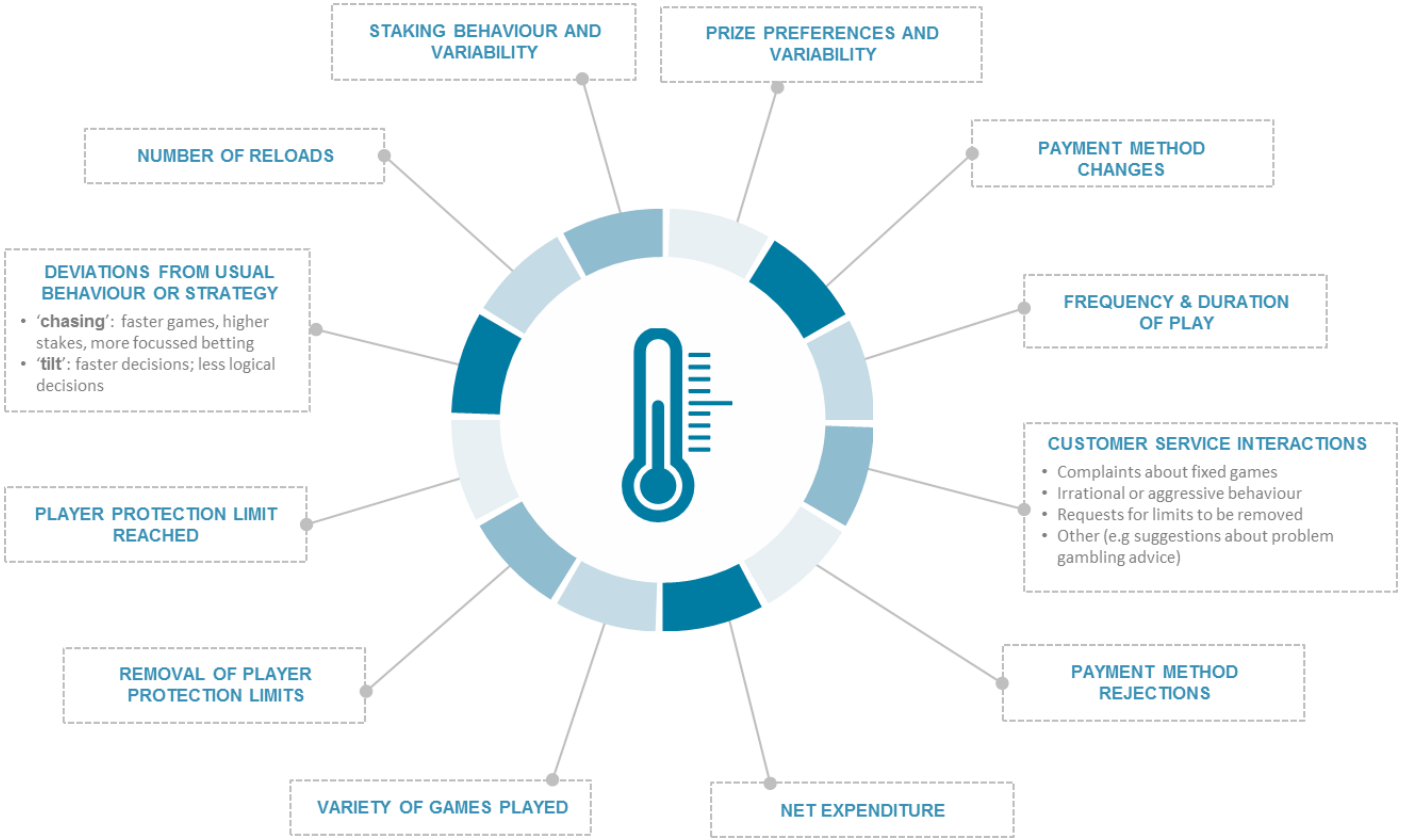
WON money:

- Bet higher stakes and withdraw more often

OWN money:

- Load more often and spend more as a percentage of their current balance

Theoretical markers of harm



Plausible ‘markers of harm’

Between Session Metrics	Within Session Metrics
1. Frequency of Play	1. Debit Card Payment Reloading and Switching
2. Duration of Play	2. Debit Card Payment Decline
3. Net Expenditure	3. Variability In Staking Behaviour
4. Levels of Play Engagement	4. Use of Autoplay
5. Number of Activities/Games Types Undertaken	5. Playing Multiple Machines Simultaneously
6. Chasing (*)	6. Stake Size
	7. Game Volatility
	8. Way Game Played
	9. Cash-Out

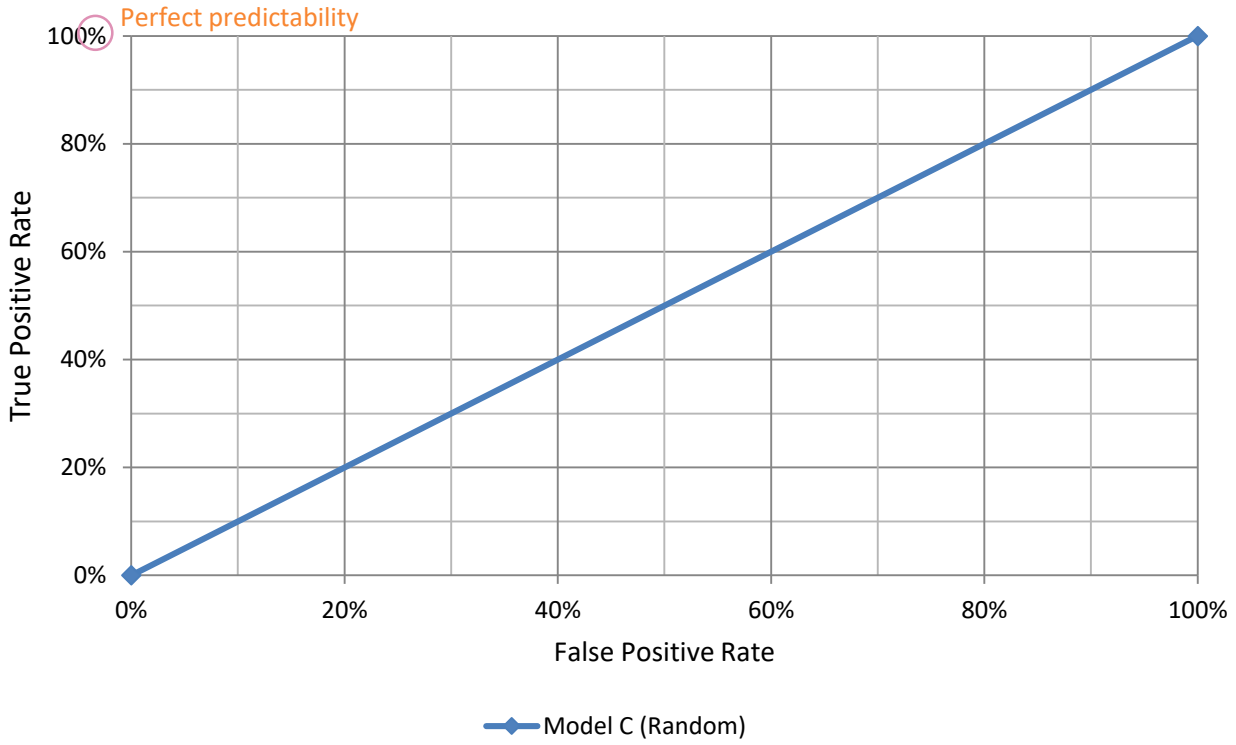
Agenda

1. Introduction
2. Data & Labels
3. Modelling
4. **Results**
5. Production deployment

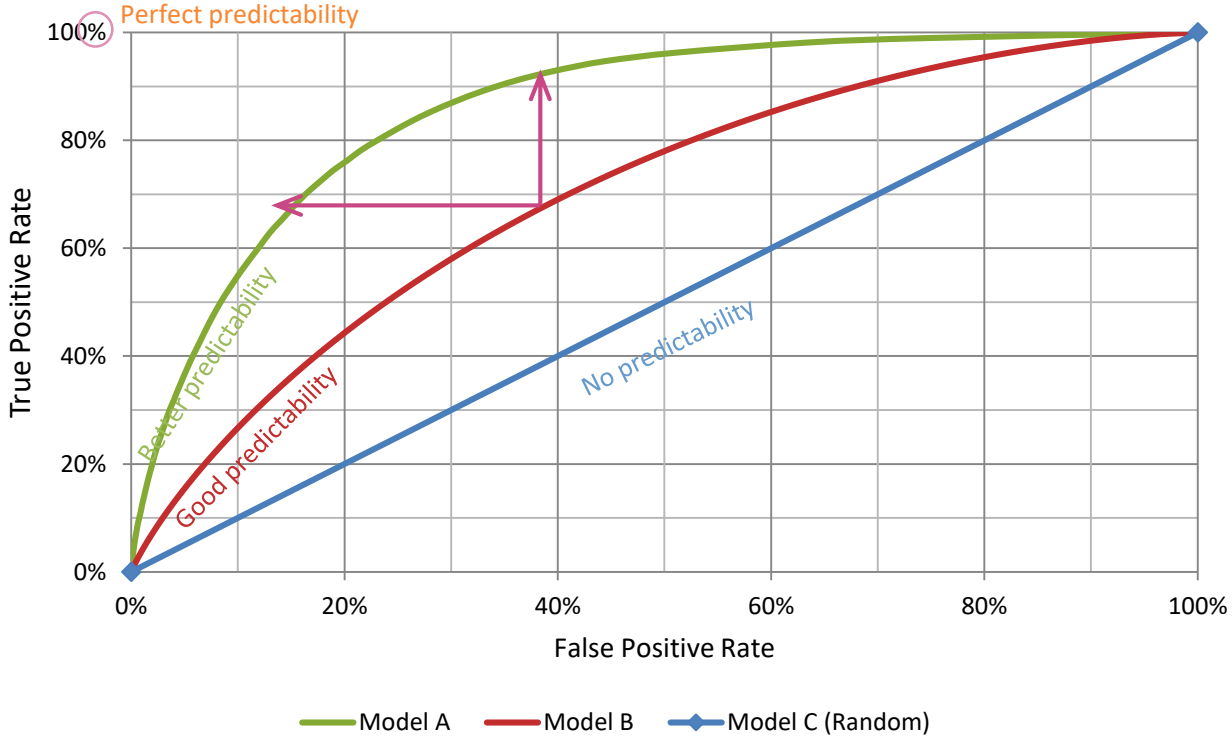
FEATURE SPACE

OUTSMART RISK

Measuring performance



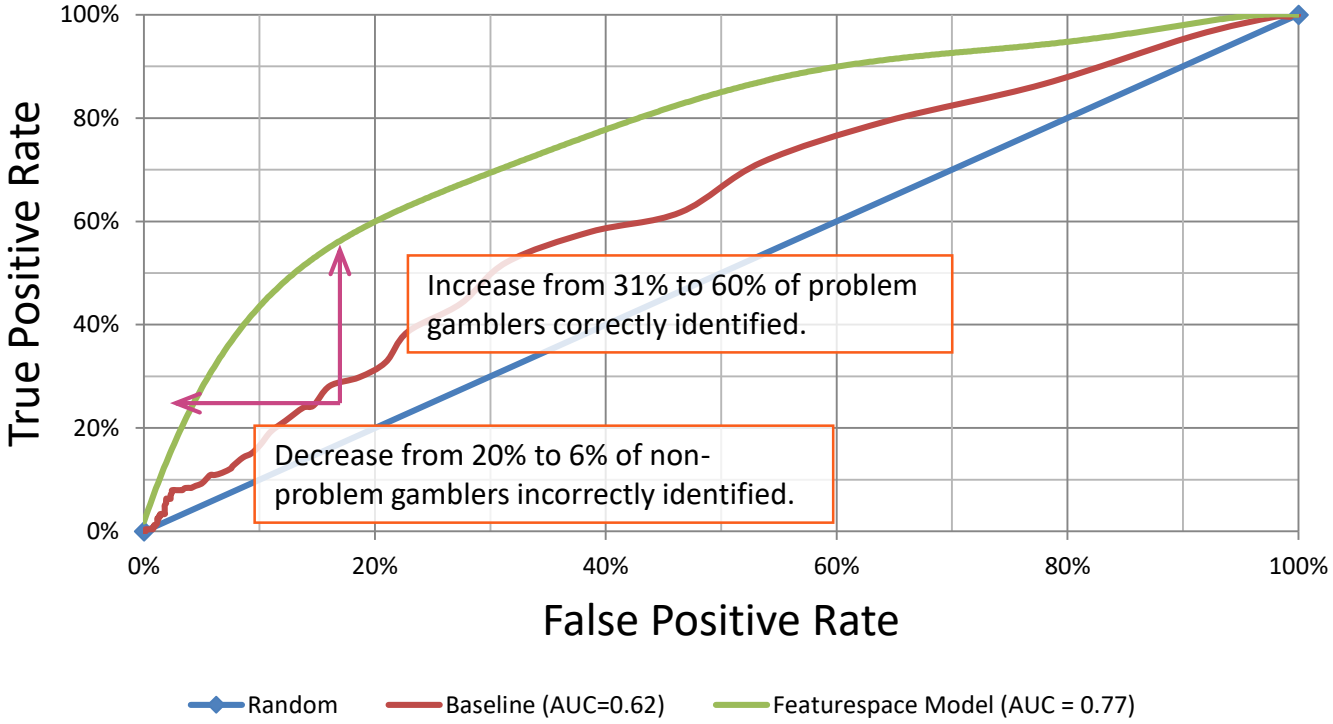
Measuring performance



Registered play

- Registered play is defined as a gaming session where a player card has been used.
- When analysing registered play, we can look at the patterns of play over multiple sessions.
- To analyse registered play:
  - All sessions from surveyed loyalty cards have been analysed.
  - A single prediction is made per loyalty card player.
  - The accuracy of the prediction is measured against the problem gambling score for that player.

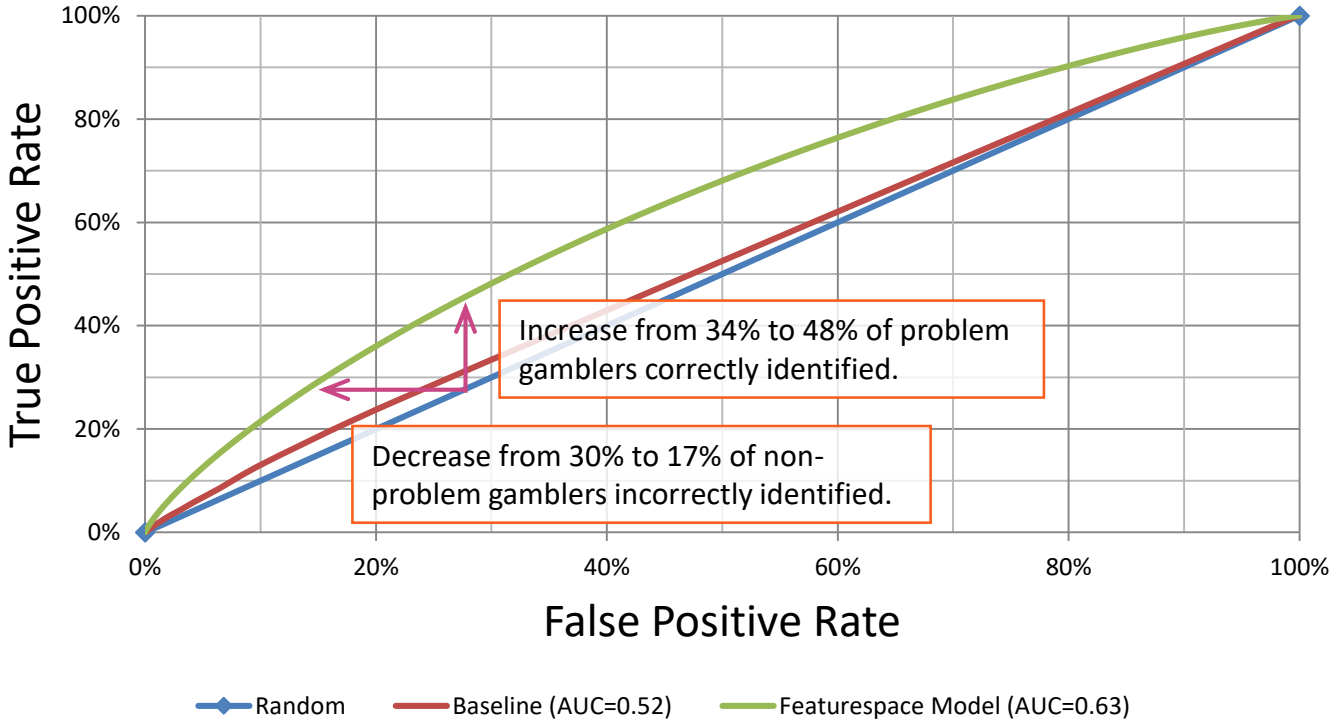
Results for registered play



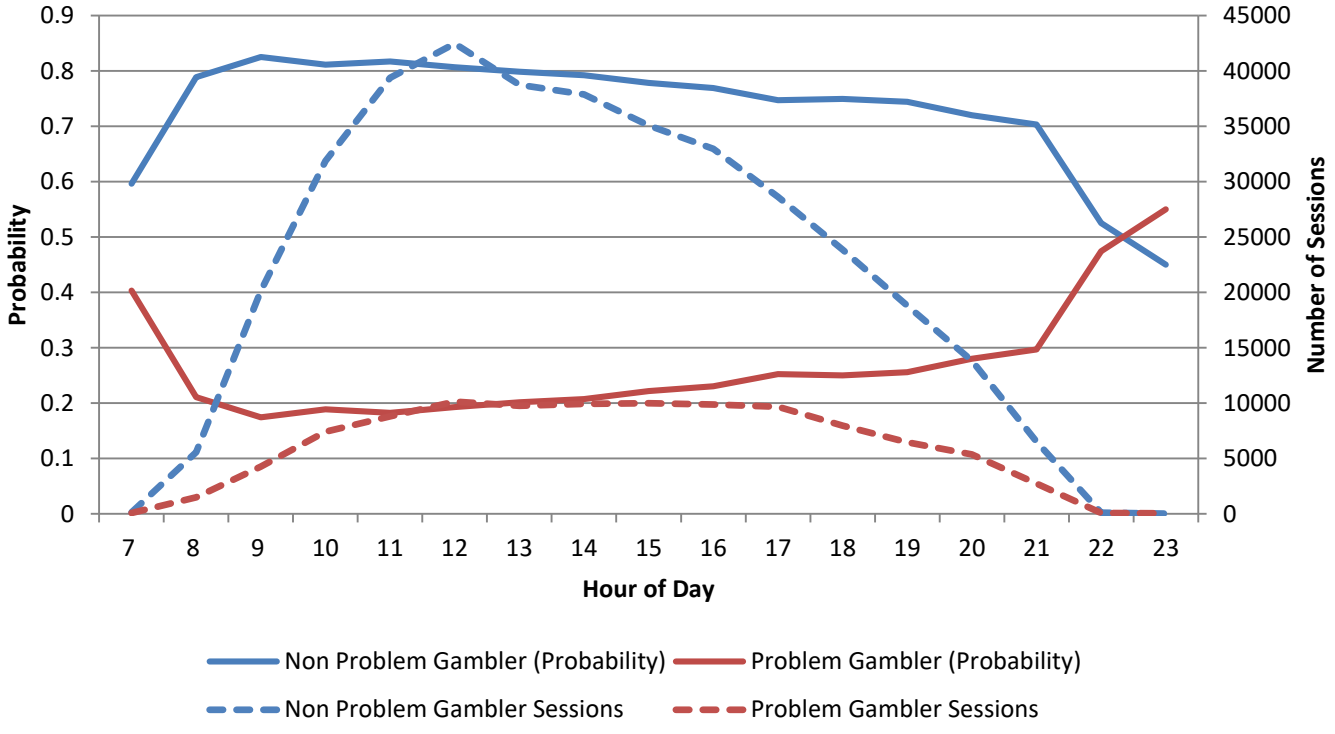
Unregistered play

- Unregistered play is defined as a gaming session where a player card hasn't been used.
- When analysing unregistered play, we have no prior history about the player.
- To analyse unregistered play:
  - All sessions from surveyed loyalty cards have been analysed.
  - A session was labelled as harmful if a problem gambler generated that session.
  - A session was labelled as non-harmful if a non-problem gambler generated that session.
  - The relationship between the session and the player is discarded.
  - A prediction is made for each session and accuracy measured accordingly.

Results for unregistered play



Time of the day



Heterogeneity among players

- Analysis results were based on ‘a’ model not necessarily ‘the’ model
- Multiple models can have similar predictive power

Model 1	Model 2
▪ Frequency of visits ▪ Variability in stake levels ▪ Hour of play ▪ Average proportion cash out	▪ Frequency of visits ▪ Game variability ▪ Total amount played in a session ▪ Difference between deposits after win and loss

- Perfect predictive model for everyone (“one model fits all”) might not be attainable, but a number of tailored models can provide a much better prediction in subgroups
- Understanding heterogeneity is important to understand who is most vulnerable
- Challenges for policy that has to work on everyone in the same way

Summary of results

- 66% improvement in accuracy of detecting problem gamblers
- Twice as many problem gamblers are correctly identified whilst maintaining a consistent false positive rate
- Additional 25% improvement in predictive accuracy for gamblers with higher PGSI scores
- Using data from a single session of play were less successful. Nonetheless, they still provided a 550% improvement on the current industry standard.

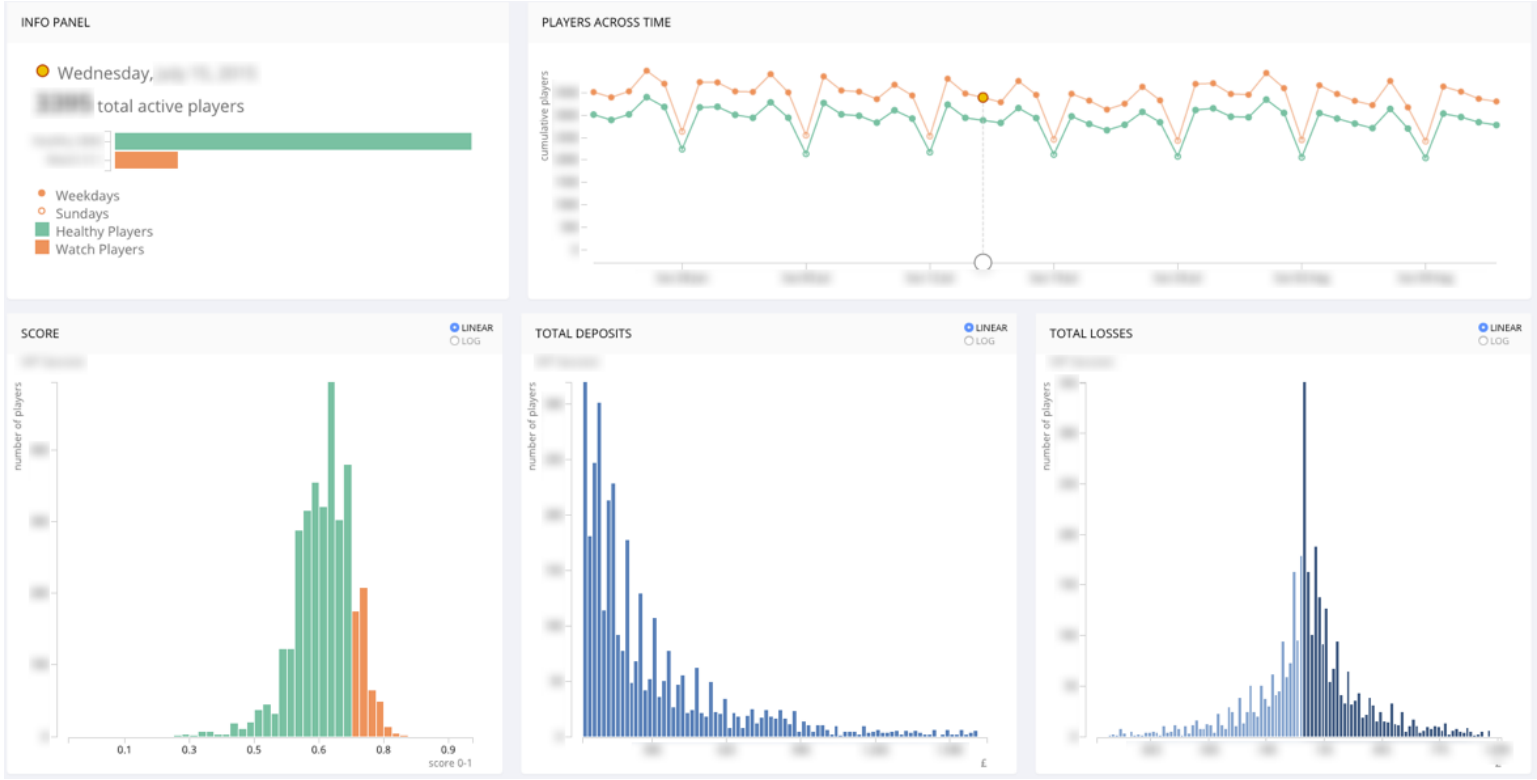
Agenda

1. Introduction
2. Data & Labels
3. Modelling
4. Results
5. **Production deployment**

FEATURE SPACE

OUTSMART RISK

Responsible Gambling Dashboard



Production deployment considerations

- Not possible to administer test to customers
  - Use proxies instead [self-exclusion/ payment failures]
- Very high value placed on model interpretability
  - Operators want to be able to justify decisions to a customer
- Feedback between interventions and training data is a very tricky problem
  - Training data available after deployment is biased

Summary

It is possible to distinguish between harmful and non-harmful gaming machine behaviour.

Furthermore,

1. It is possible to score individual players and sessions based on a harm-related risk score.
    - 66% improvement in accuracy of detecting problem gamblers
    - These players can be added to a watch list or receive targeted interventions.
2. Gambling behaviours are complex. Identifying gambling related harm is complex.
    - There isn't a simple criteria that can be used to identify this behaviour.
    - By applying predictive behavioural technology, a solution can be operationalised.


FEATURE
SPACE
OUTSMART RISK