

Cash Analytics

Eduardo Contreras Cortes

EY UK Data Analytics

A collage of various international banknotes, including Euro, Indian Rupee, Vietnamese Dong, and others, scattered across the frame. The text "How do you predict payments?" is overlaid in the center.

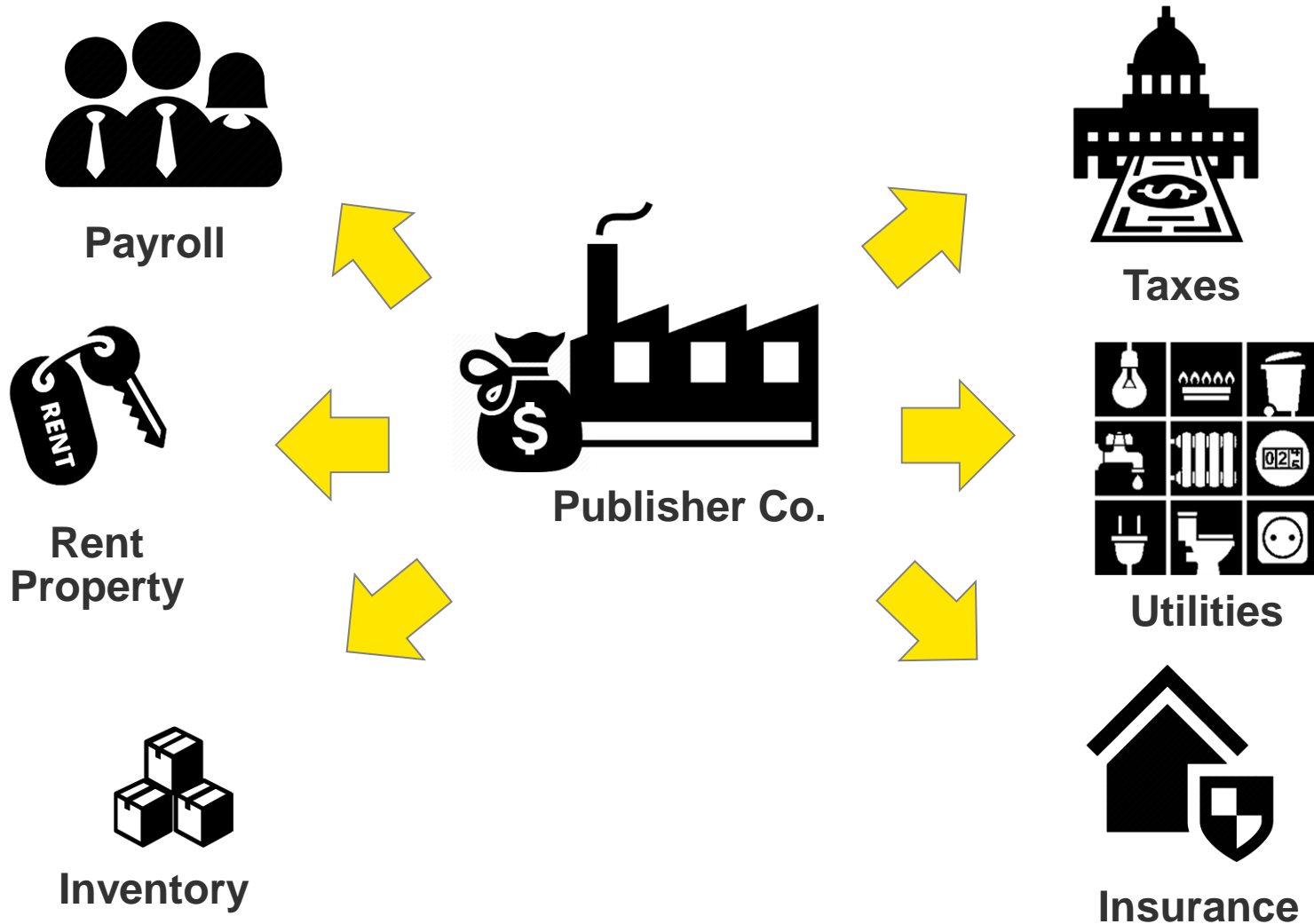
How do you predict payments?

29%

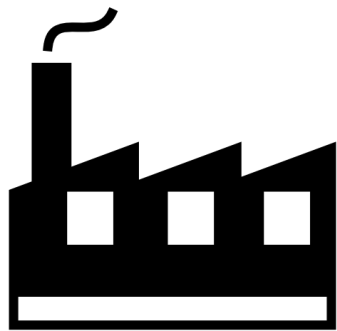
Share of commercial payments in the UK
not paid on time



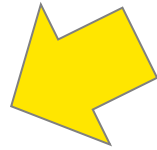
Cash is critical for every company



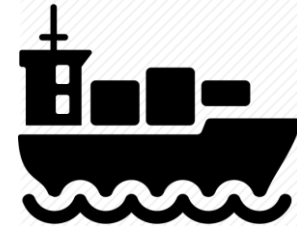
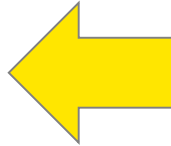
Cash trapped in late payment is cash you can't spend



Publisher Co.



Company A



Company B



Company C

Can we use analytics to predict which ones?

The Approach

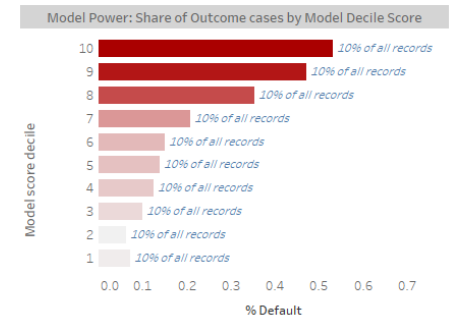
I. Data extraction and feature engineering



II. Data analysis and model ensemble







III. Model insights





I. Data Extraction and Feature Engineering

- ▶ Understand the business side
- ▶ Create a single company view. Some companies are spelled different through time (“Ernst & Young”, ”Ernstt & Young”, “Ernst Young”)
- ▶ Built more than 150 features:

Example of type of features	
	Type of company (location, sector, size)
	Early/Late payments frequency
	Early/Late payments amounts
	Timing of payments (month, day of week)



II. Data analysis and model ensemble

We used several R libraries, a huge time saver!

1

Descriptive Analysis

- **Smbinning** great for summarising descriptive statistics (uniques, zeroes, % Missing) and initial predictive profiling

2

Correlation Analysis

- **Corr & Vif functions** To remove highly correlated variables and detect how variables interact

3

Additional Feature Engineering

- **Interactions** try two-way or three-way to see if new features provide better predictions
- **Reclassification of categorical variables** some algorithms only allow numerical variables

4

Model Ensemble

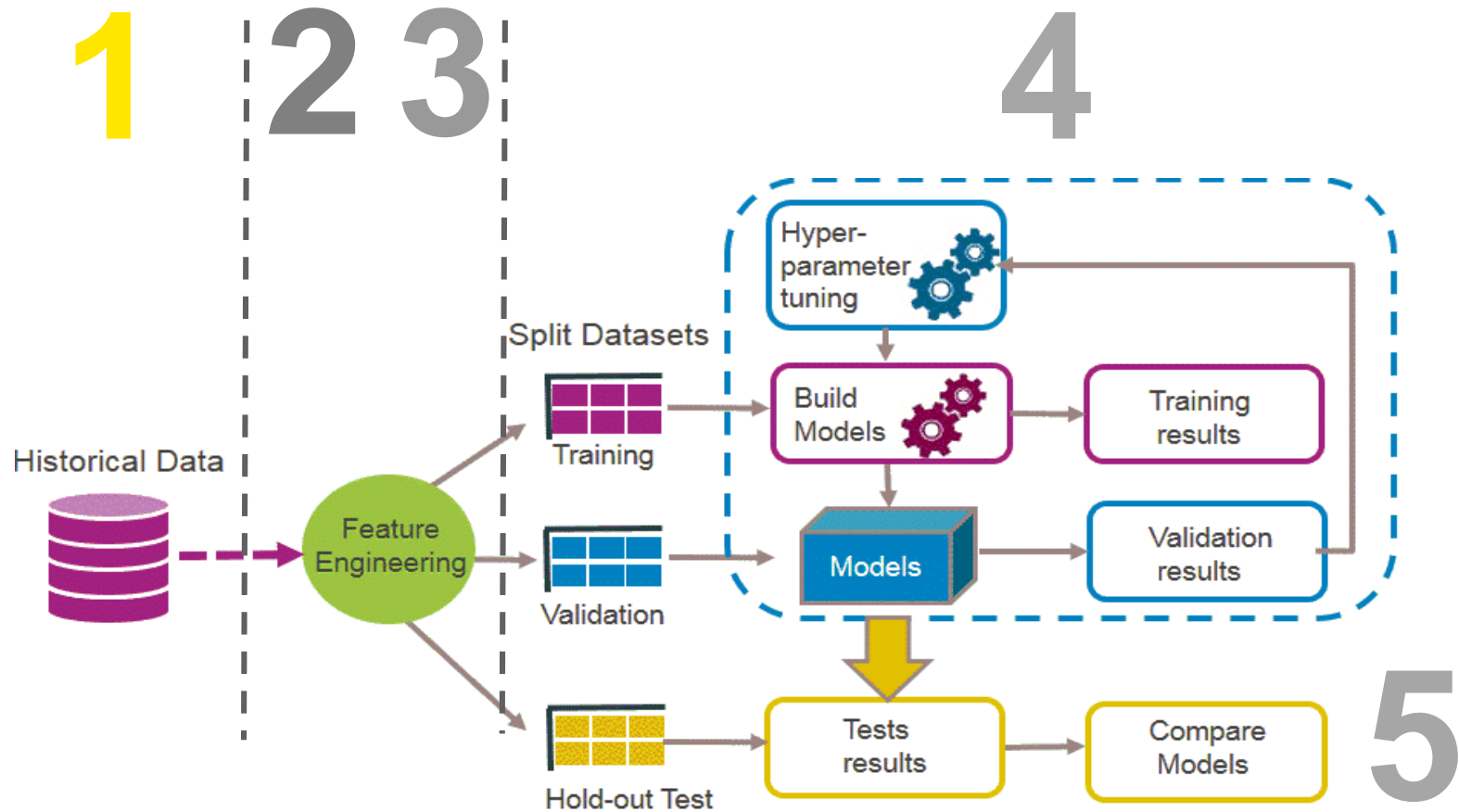
- **Lm** start with a simple regression model
- **Ranger** a fast implementation of random forest
- **Xgboost** a fast implementation of gradient boosting methodology

5

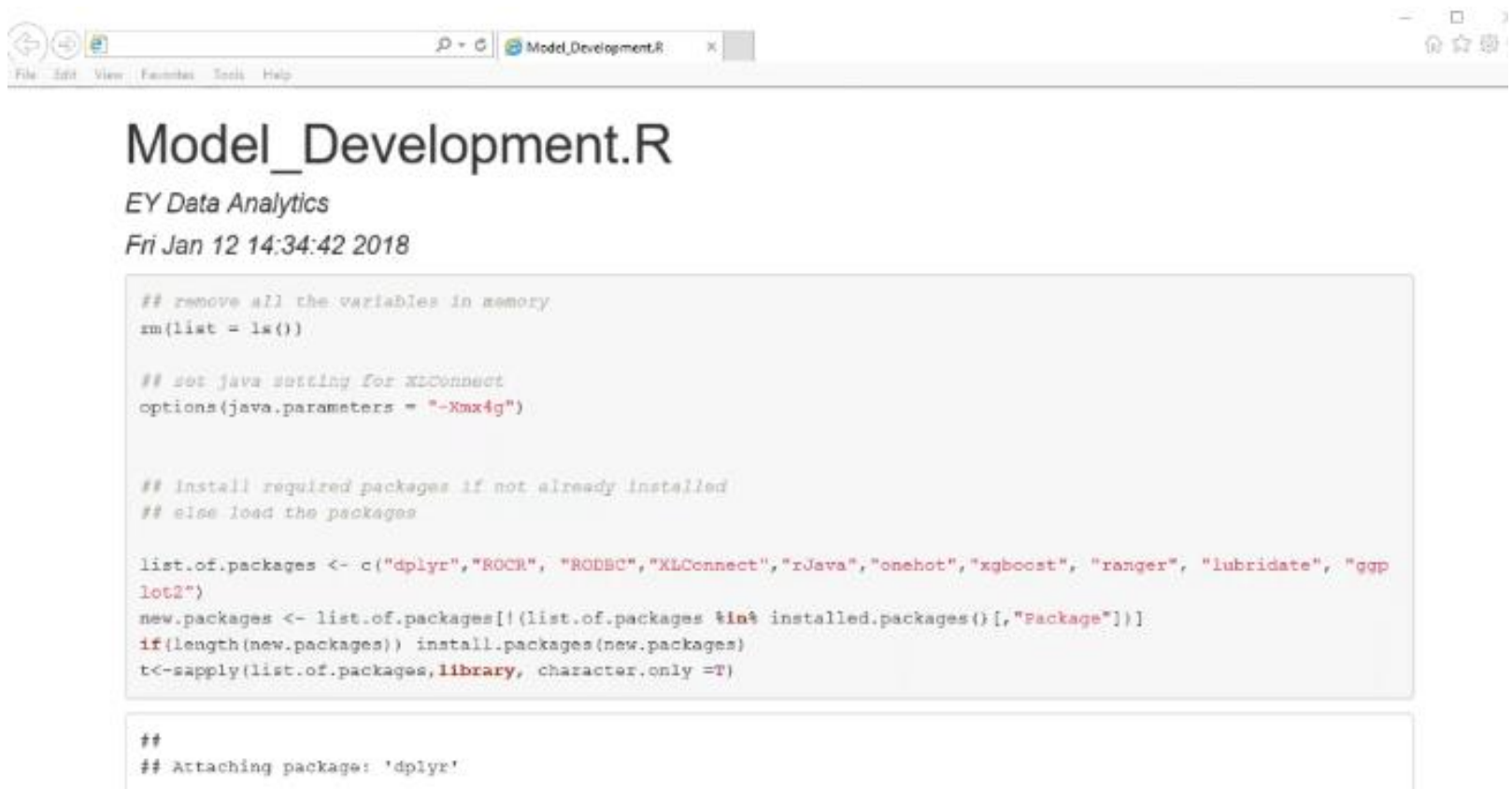
Model Assessment

- **From simplex to complex models**
- **Understand the most relevant variables of the models**
- **Average models from different techniques** to average disadvantages of each technique to reduce errors

R II. Model ensemble



R II. Model ensemble



```
## remove all the variables in memory
rm(list = ls())

## set java setting for XLConnect
options(java.parameters = "-Xmx4g")

## Install required packages if not already installed
## else load the packages

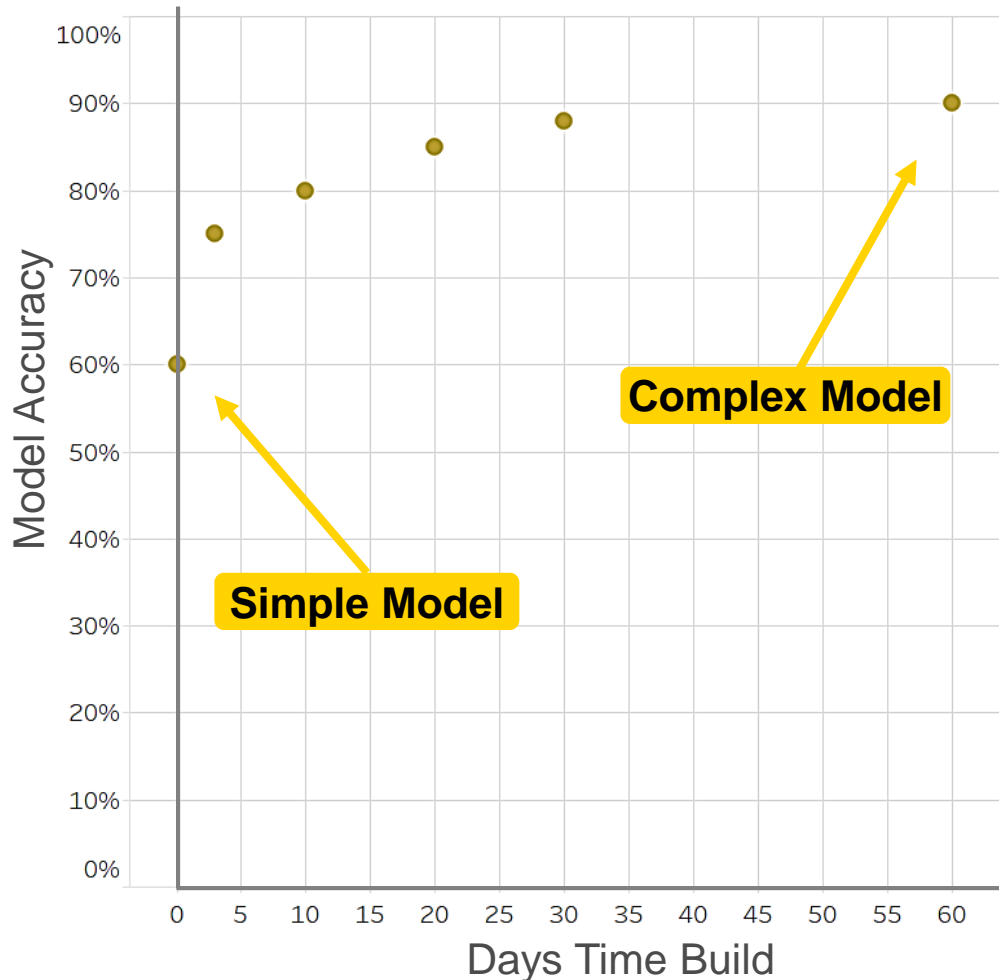
list.of.packages <- c("dplyr", "ROCR", "RODBC", "XLConnect", "rJava", "onehot", "xgboost", "ranger", "lubridate", "ggplot2")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages)
t<-sapply(list.of.packages, library, character.only =T)

##
## Attaching package: 'dplyr'
```

**Keep track of model parameters and results
with R Markdown**

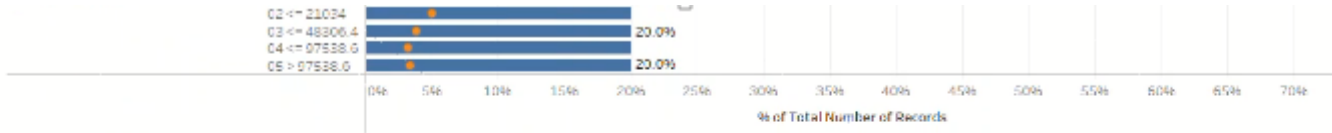
R II. Model ensemble

Model Accuracy vs Days Time Build



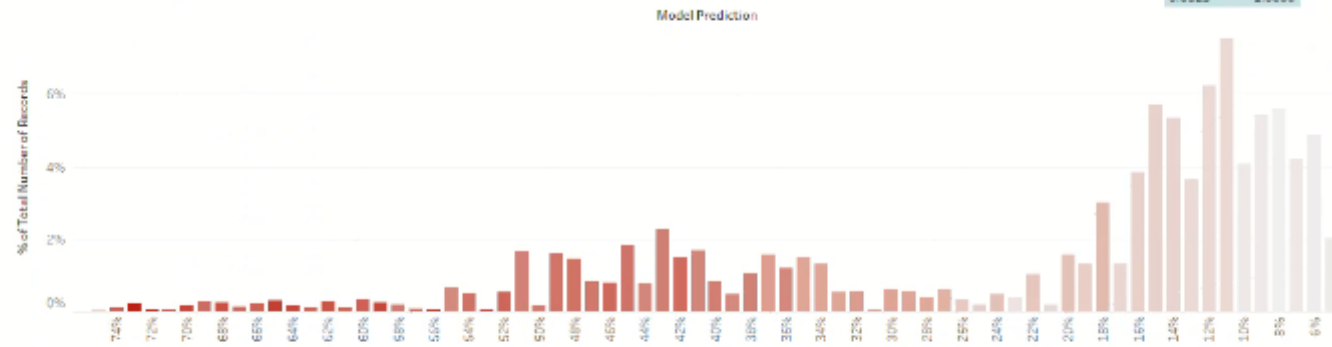
30% uplift was
£2.5M of late
payments
identified for
Publisher Co.

III. Model insights



Percentage of observations by Model Probability

Model Distribution



Detailed Data

Predict Bin	ID	Outcome	Prediction	Variables				
				AVE_UTIL_group	LIMIT_BAL_group	MAX_BILL_group	MAX_PAY_group	RATIO_PAY_BILL_group
1	1	0	5%	02 <= 0.134	05 > 270000	04 <= 97538.6	01 <= 0	05 > 0.767
	2	0	5%	02 <= 0.134	06 > 270000	04 <= 97538.6	01 <= 0	05 > 0.767

III. Model insights



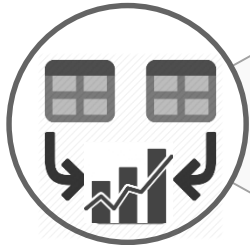
In summary



Machine Learning helped **to identify** companies with high risk of late payments



Companies were able **to collect cash** that was trapped in late payments



Companies can analyse the business impact by creating different scenarios using **interactive reports**



Ability to **share** with minimal effort the business benefits of the models with different audiences

**To find more about EY Cash Analytics
visit:**

<http://bit.ly/EYCashAnalyticsVideo>

Contact me at:



Eduardo Contreras Cortes

www.linkedin.com/in/eduardocontrerasc

EY | Assurance | Tax | Transactions | Advisory

Ernst & Young LLP

© 2015 Ernst & Young LLP. Published in the UK.
All Rights Reserved.

The UK firm Ernst & Young LLP is a limited liability partnership registered in England and Wales with registered number OC300001 and is a member firm of Ernst & Young Global Limited.

Ernst & Young LLP, 1 More London Place, London, SE1 2AF.

ey.com