WENDY WONG,
masteR of Data science student at UTs

FROM FINANCIAL PLANNNG TO DATA SCIENCE

- FINANCIAL PLANNING AT WESTPAC, FUNDS MANAGEMENT AT MACQUARIE, ADVICE COMPLAINTS AT AMP

Aviation career

be resilient.
sometimes you crash

GFC
(AFTER FINANCIAL PLANNING)

What is your Ikigai?

World flight in 2014

INSTITUTE OF ANALYTIC PROFESSIONALS AUSTRALIA – EMBRACE THE UNKNOWN

Speculation on disappearance

Data Science is an ART.
Be creative.
FIND an interesting dataset from Kaggle, UCI Machine Learning Repository
Practice on the
eg. Wine set, boston housing

GOODBYE TO FINANCIAL PLANNING & HELLO ANALYTICS

Working in Agile projects at ATo

practice with sample data
explore the data
Important to pre-process the data using
dplyr
what type of problem?
logistics regression for classification,
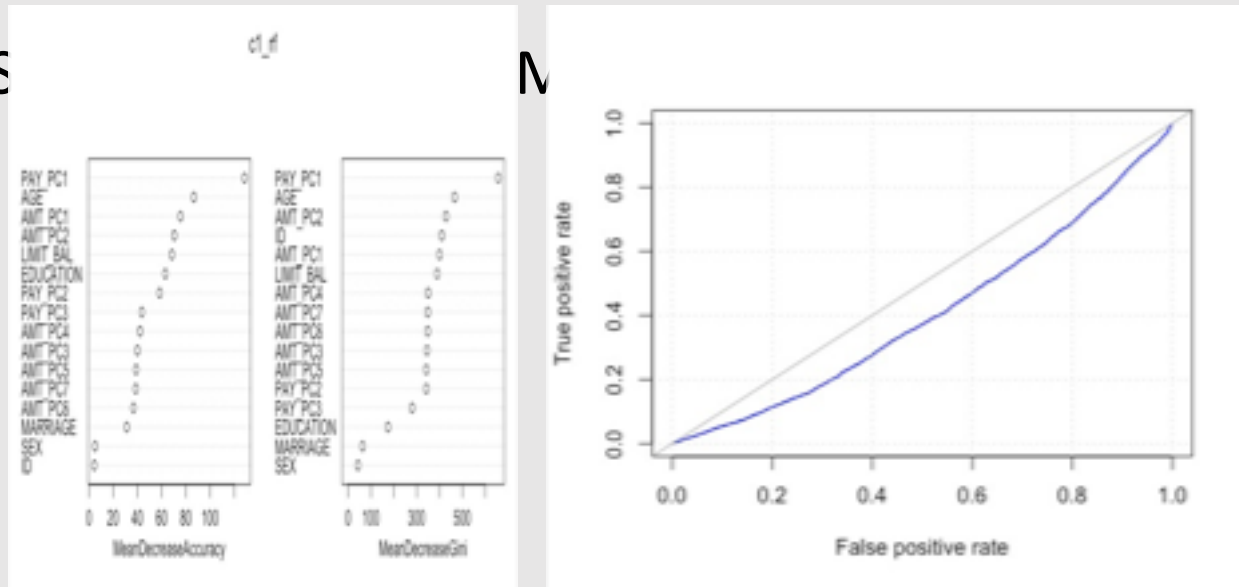decision tree, random forest, Xgboost for
fast computation
understand what stakeholder wants
embrace uncertainty
TRY to FIX debugging

Packages used for machine learning at ATO

- DPLYR – DATA WRANGLING

- DUMMIES – ONE HOT ENCODING

- CARET – DATA PARTITIONING

- GGPLOT2 – EDA

- LOGISTIC REGRESSION – GLM()

- GBM – GRADIENT BOOSTING

- CTREE () DECISION TREE – SMALL BUSINESS TAX ON-TIME

- RANDOMFOREST– FEATURE IMPORTANCE

Books

- The Art of Data Science by Roger Peng

- Exploratory Data Analysis Using R by Roger Peng

- R Programming for Data Science by Roger Peng

- An Introduction to Statistical Learning with applications using R by Robert Tibshirani et al

WEBSITES

- Towards Data Science

- KDNuggets

- R Bloggers

- Data Camp cheat sheets

THANK YOU

@abc_wendsss