Making Multimillion-Dollar Decisions with H_2O AutoML, LIME and Shiny

Jo-fai (Joe) Chow

Data Science Evangelist / Community Manager

joe@h2o.ai

@matlabulous

Download → https://bit.ly/

h2o_meetups

About this talk

- Quick Overview
  - Business problem
  - Solution and result
- Details
  - The “Moneyball” team
  - Baseball data → ML problem
  - H_2O AutoML, LIME, Shiny

You will learn …

- How to frame a business problem (e.g. Moneyball) for machine learning.
- If we have time …
  - How to use H_2O AutoML with R interface.
  - How to use LIME to explain H_2O models.

About Me

- **Before H_2O**
  - Water Engineer / EngD Researcher / Matlab Fan Boy
    (wonder why @matlabulous?)
  - Discovered R, Python, H_2O … never look back again
  - Data Scientist at Virgin Media (UK), Domino Data Lab (US)
- **At H_2O …**
  - Data Scientist / Evangelist /
  - Sales Engineer / Solution Architect /
  - Community Manager

    … The harsh reality of startup life …
  - H_2O SWAG Photographer

    #AroundTheWorldWithH2Oai

    Love H_2O? Get some stickers!

Jo-fai (Joe) Chow
@matlabulous

Thanks all for coming to my @erum2018 workshop. Here is our #360selfie. Hope you all enjoyed building @h2oai models w/ #AutoML and explaining them w/ #LIME. Looking forward to the welcome reception and #Shiny demos - totally my thing! #eRum2018 #Budapest #AroundTheWorldWithH2Oai

4:45 PM - 14 May 2018 from Budapest, Hungary

Jo-fai (Joe) Chow
@matlabulous

Another #FullHouse @h2oai #LondonAI #meetup tonight. Thanks @MSFTReactor for the amazing venue and food! #OpenSource #Community #MVPBuzz #AroundTheWorldWithH2Oai #360Selfie 🇬🇧 cc our guest speakers @SKREDDY99 @cheukting_ho & Josh Warwick

7:15 PM - 12 Mar 2018 from London, England

Jo-fai (Joe) Chow
@matlabulous

Awesome #KNIMESummit2018 #KNIMESpringSummit in #Berlin. @knime @Kuriooos Marten here is our #360Selfie cc @h2oai #AroundTheWorldWithH2Oai 🇩🇪 #OpenSource #MachineLearning #Community 💪

1:54 PM - 7 Mar 2018 from Hotel Berlin

Jo-fai (Joe) Chow
@matlabulous

Thanks @ingnl for hosting @h2oai #meetup in #Amsterdam last week. Tremendous turnout and great discussions. #AroundTheWorldWithH2Oai #360Selfie 🇳🇱 cc @fishnets88

7:15 AM - 26 Feb 2018 from Amsterdam, The Netherlands

Jo-fai (Joe) Chow
@matlabulous

Merci beaucoup Alexia, Samia & Aurelie from @tlse_dasci. We had our very first @h2oai #meetup in #Toulouse tonight. Fantastic crowd and awesome @HarryCoworking venue. We hope to see you all again in the future. Here is our #360selfie 📸 #AroundTheWorldWithH2Oai 🇫🇷

10:35 PM - 23 Apr 2018 from Toulouse, France

Jo-fai (Joe) Chow
@matlabulous

@h2oai #DeepWater at #KoelnRUG #meetup thanks for having us. Slides: github.com/h2oai/h2o-meet ... #AroundTheWorldWithH2Oai #Cologne #Germany 🇩🇪

6:34 PM - 17 Mar 2017 from Cologne, Germany

Reminder: #360Selfie

H2O.ai

About **H2O.ai** …

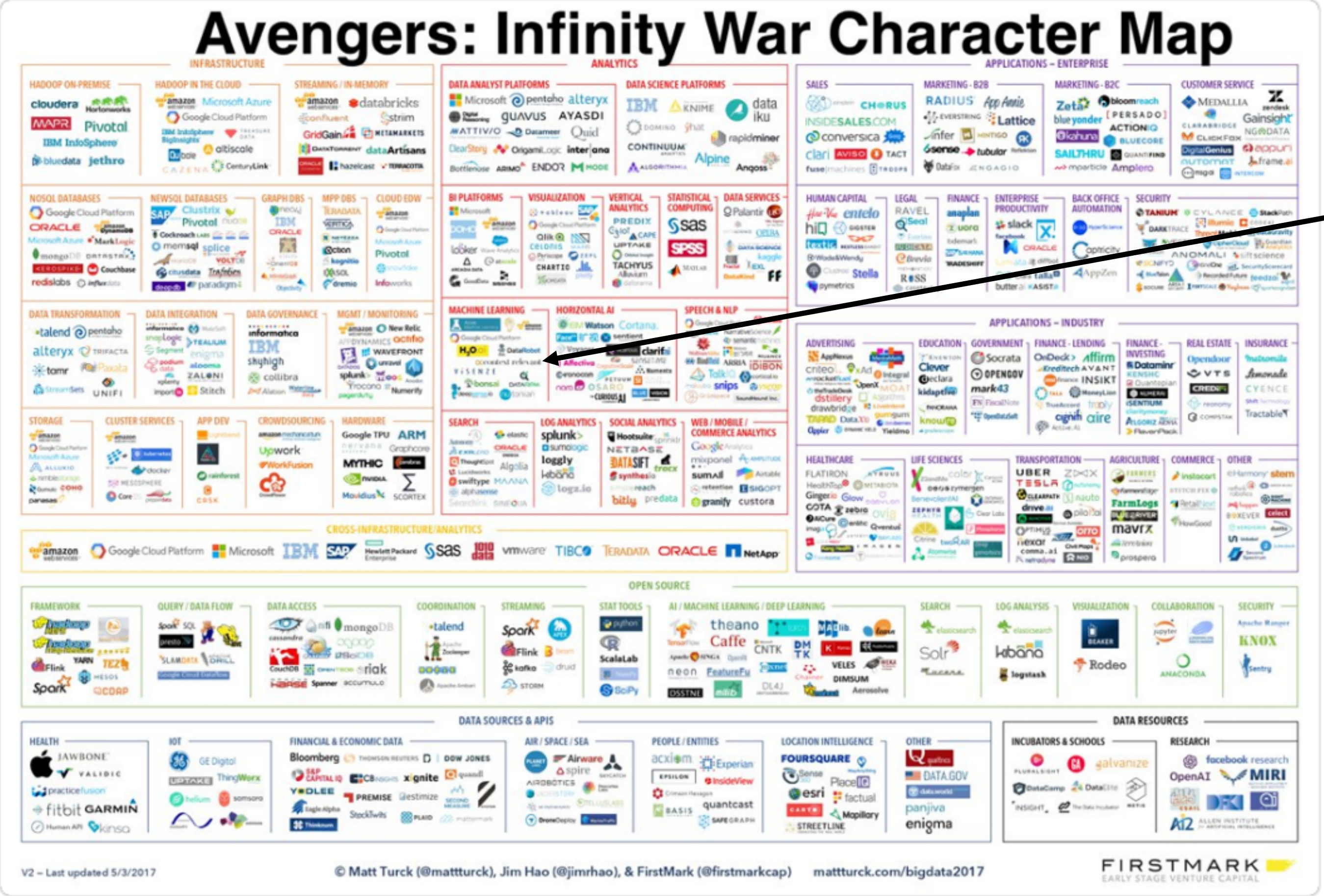
Have you seen Avengers: Infinity War?

Do you know all the characters in the movie? (No spoilers - I promise)

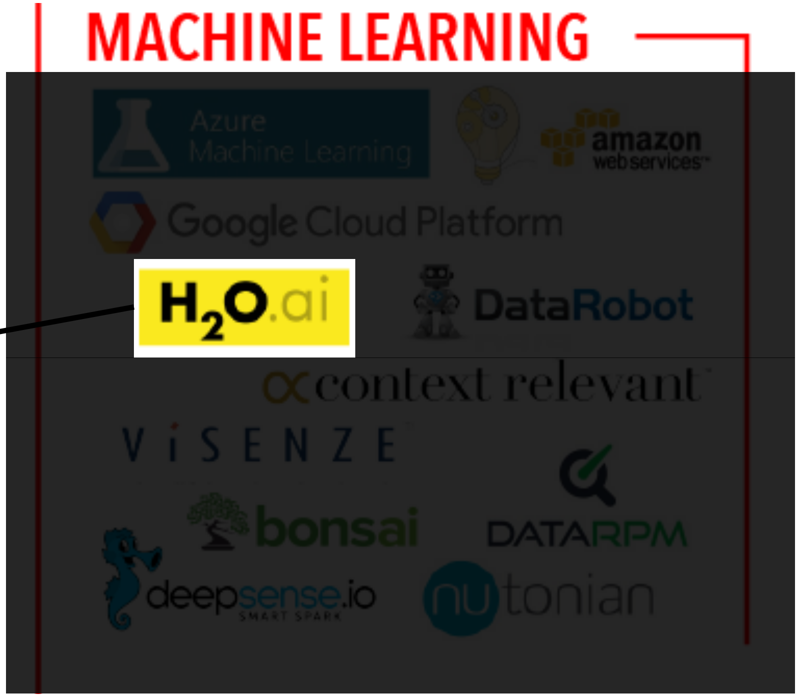
Vicki Boykis
@vboykis

Follow

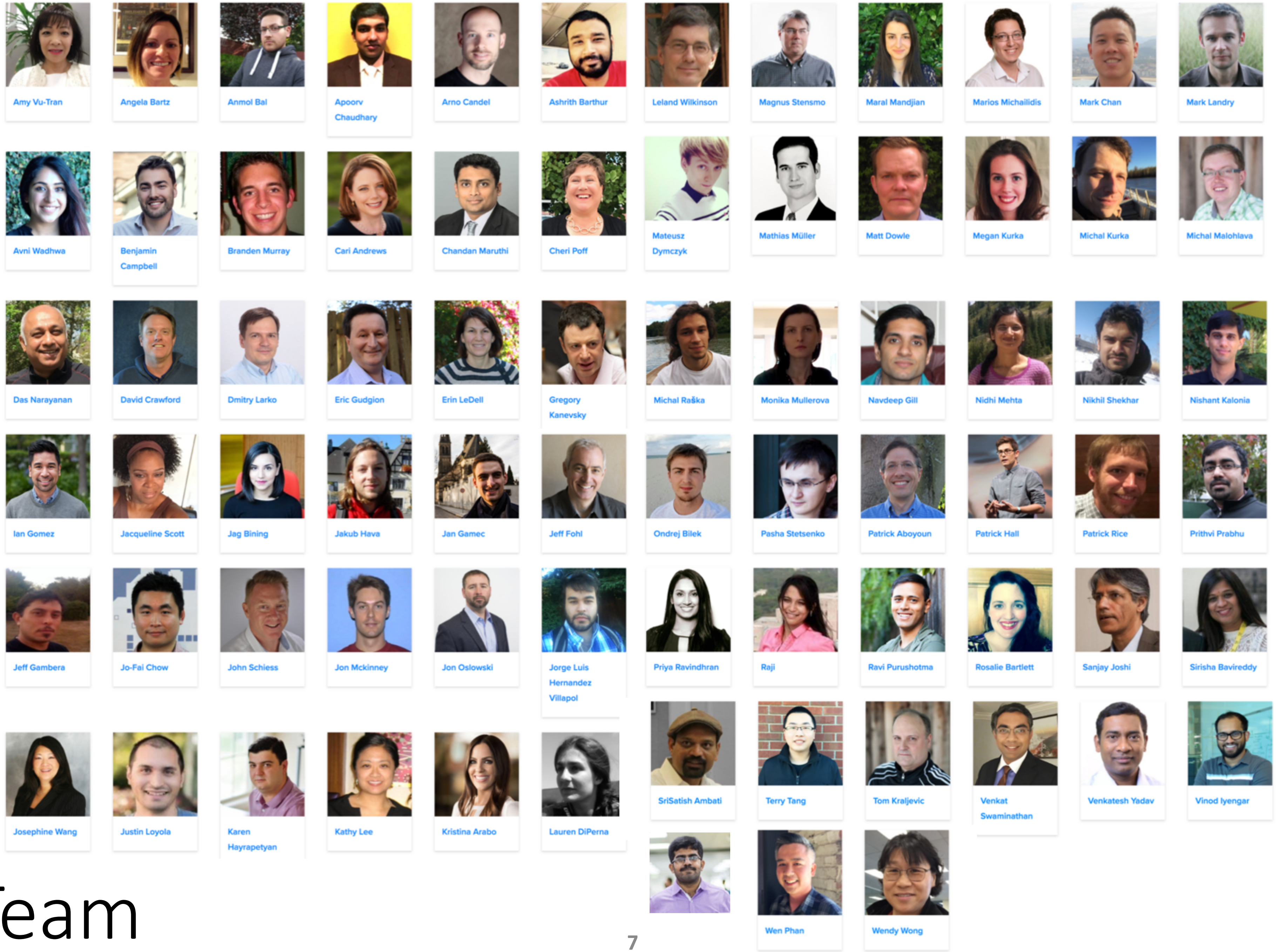
I made a guide for anyone who was as confused by all the characters in Infinity War as I was.



1:06 AM - 11 May 2018



We develop
machine learning platforms


Amy Vu-Tran
Angela Bartz
Anmol Bal
Apoorv Chaudhary
Arno Candel
Ashrith Barthur
Leland Wilkinson
Magnus Stensmo
Maral Mandjian
Marios Michailidis
Mark Chan
Mark Landry
Avni Wadhwa
Benjamin Campbell
Branden Murray
Cari Andrews
Chandan Maruthi
Cheri Poff
Mateusz Dymczyk
Mathias Müller
Matt Dowle
Megan Kurka
Michal Kurka
Michal Malohlava
Das Narayanan
David Crawford
Dmitry Larko
Eric Gudgion
Erin LeDell
Gregory Kanevsky
Michal Raška
Monika Mullerova
Navdeep Gill
Nidhi Mehta
Nikhil Shekhar
Nishant Kalonia
Ian Gomez
Jacqueline Scott
Jag Bining
Jakub Hava
Jan Gamec
Jeff Fohl
Ondrej Bilek
Pasha Stetsenko
Patrick Aboyoun
Patrick Hall
Patrick Rice
Prithvi Prabhu
Jeff Gambera
Jo-Fai Chow
John Schiess
Jon Mckinney
Jon Oslowski
Jorge Luis Hernandez Villapol
Priya Ravindhran
Raji
Ravi Purushotma
Rosalie Bartlett
Sanjay Joshi
Sirisha Bavireddy
Josephine Wang
Justin Loyola
Karen Hayrapetyan
Kathy Lee
Kristina Arabo
Lauren DiPerna
SriSatish Ambati
Terry Tang
Tom Kraljevic
Venkat Swaminathan
Venkatesh Yadav
Vinod Iyengar
Wen Phan
Wendy Wong

H_2O Team

Gartner names H2O as **Leader** with the **most completeness of vision**

- H2O.ai recognized as a **technology leader with most completeness of vision**

- H2O.ai was recognized for the mindshare, partner network and status as a **quasi-industry standard** for machine learning and AI.

- **H2O customers gave the highest overall score** among all the vendors for sales relationship and account management, customer support (onboarding, troubleshooting, etc.) and overall service and support.

Figure 1. Magic Quadrant for Data Science and Machine-Learning Platforms

Source: Gartner (February 2018)

Platforms with H_2O integration

srisatish
@srisatish

Following

Replying to @BobMuenchen @knime @h2oai

@KNIME gained the ability to run @H2O.ai algorithms, so these two may be viewed as complementary, not competitors #Ecosystem #OpenSource

3:32 PM - 2 Mar 2018

H_2O + KNIME Talk at KNIME Summit Mar 2017

1:54 PM - 7 Mar 2018 from Hotel Berlin

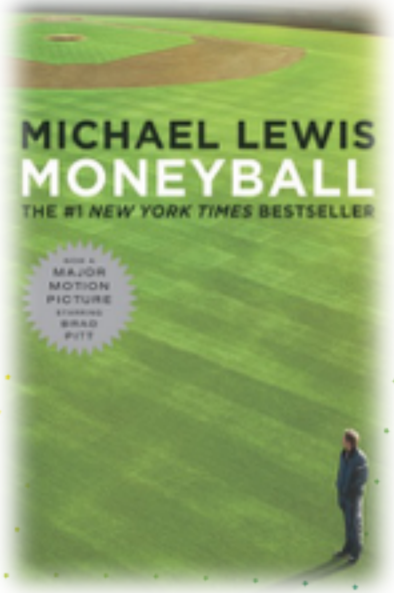
Figure 1. Magic Quadrant for Data Science and Machine-Learning Platforms

Source: Gartner (February 2018)

Moneyball:
The Multimillion-Dollar Business Problem

The quest to find the most undervalued baseball players (before other teams notice them)

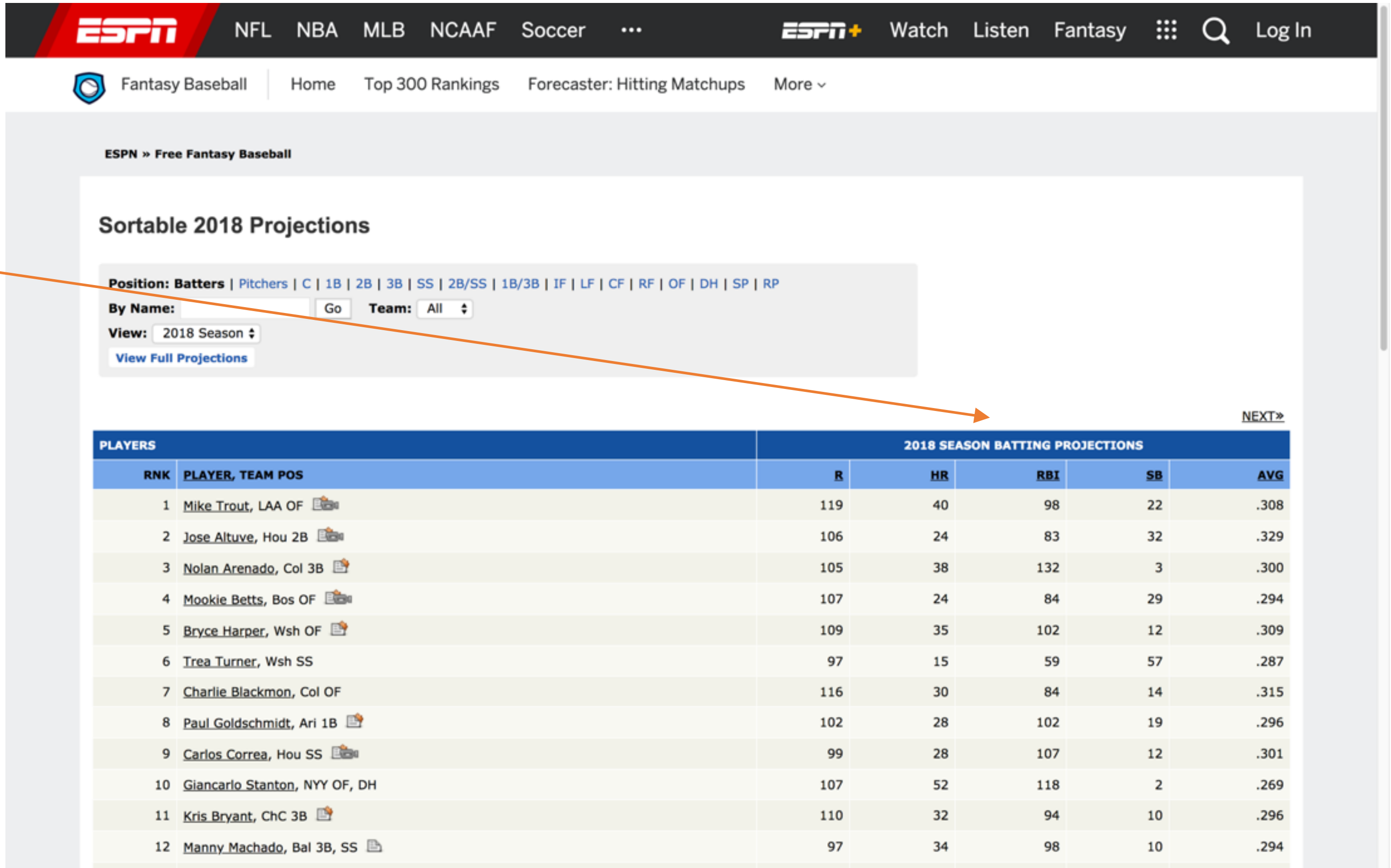
Source: Moneyball, 2011 Columbia Pictures



H2O.ai

The Real Business Problem in Major League Baseball (MLB)

- Existing Forecasts (e.g. ESPN) are usually projections for the next year only.
- MLB players usually consider terms for 3 to 5 years when they sign a new contract.
- MLB teams need to consider players' long-term performance (i.e. > 1 year)



PLAYERS		2018 SEASON BATTING PROJECTIONS				
RNK	PLAYER, TEAM POS	R	HR	RBI	SB	AVG
1	Mike Trout, LAA OF	119	40	98	22	.308
2	Jose Altuve, Hou 2B	106	24	83	32	.329
3	Nolan Arenado, Col 3B	105	38	132	3	.300
4	Mookie Betts, Bos OF	107	24	84	29	.294
5	Bryce Harper, Wsh OF	109	35	102	12	.309
6	Trea Turner, Wsh SS	97	15	59	57	.287
7	Charlie Blackmon, Col OF	116	30	84	14	.315
8	Paul Goldschmidt, Ari 1B	102	28	102	19	.296
9	Carlos Correa, Hou SS	99	28	107	12	.301
10	Giancarlo Stanton, NYY OF, DH	107	52	118	2	.269
11	Kris Bryant, ChC 3B	110	32	94	10	.296
12	Manny Machado, Bal 3B, SS	97	34	98	10	.294

H2O.ai

Our Solution

- Open data – Lahman Database.
- Proprietary data from Ari Kaplan – our real Moneyball guy.
- Framed data for ML.
- Used H_2O AutoML to make predictions for next three years.
- Created a Shiny app for quick navigation.
- Ari used the app to look at predictions for some free agents. He found an undervalued player and recommended that player to his team.
- The rest is history.

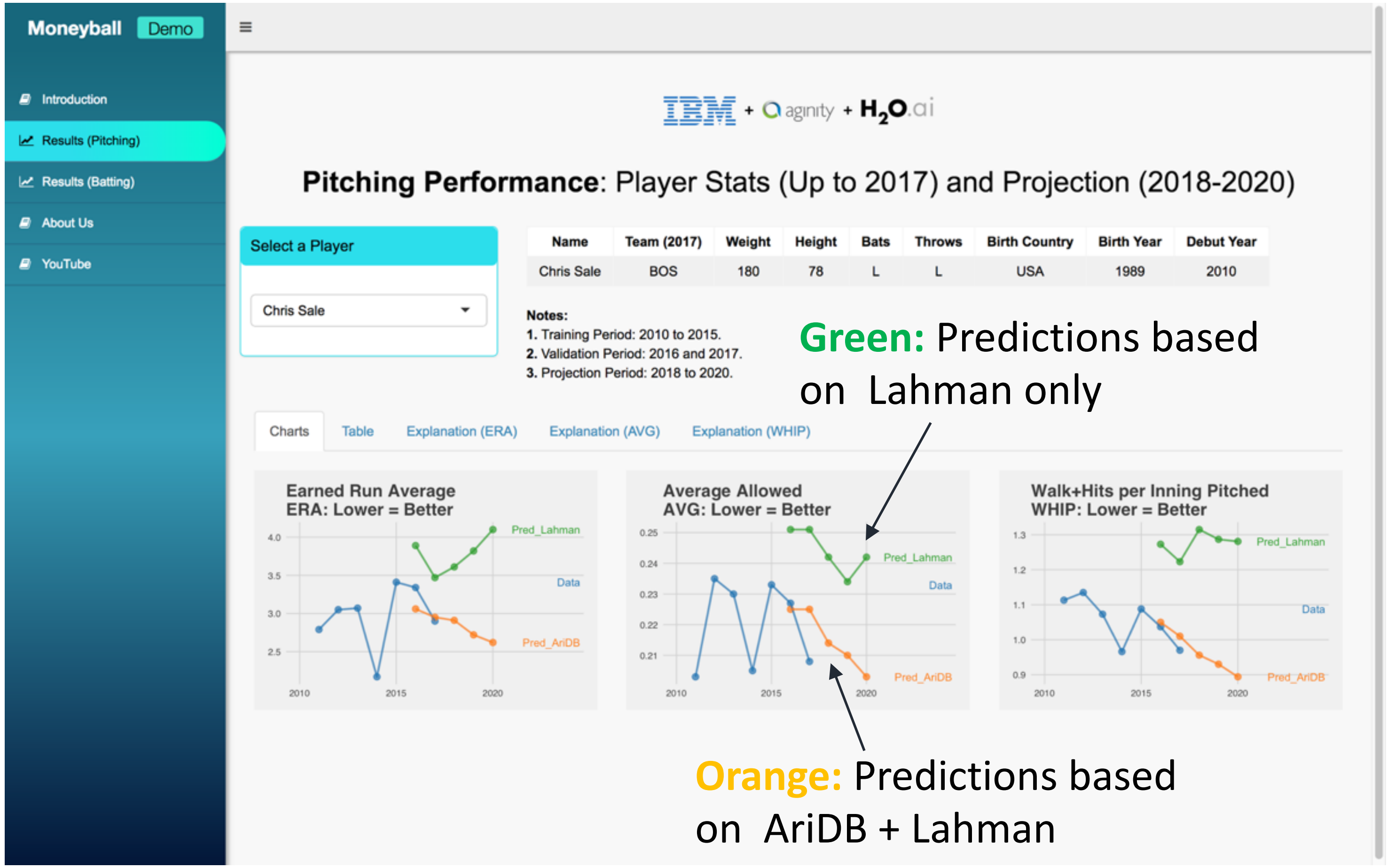
In case you're wondering… final project result

$20M

trade 2 weeks prior to the season beginning

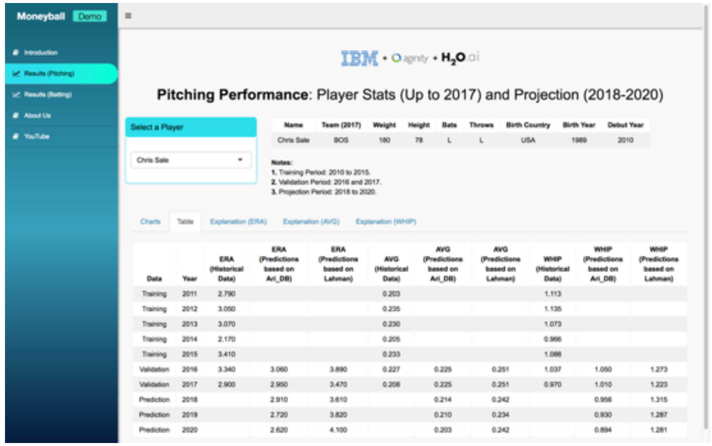
IBM + aginity + H_2O.ai

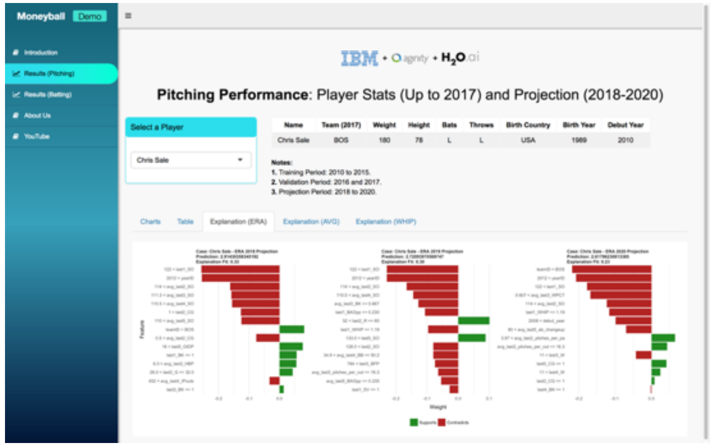
Live Demo



Green: Predictions based on Lahman only

Orange: Predictions based on AriDB + Lahman





More Details

- Moneyball team
- How to frame the Moneyball problem for automatic machine learning
- Tools: H_2O AutoML, LIME, Shiny …

The Moneyball Team

David Kearns

PM @ IBM Data Science

Ari Kaplan

Mr. Moneyball @ Aginity

Jo-Fai Chow

Data Scientist @ H_2O.ai

Ari Kaplan – the Real Moneyball

- The real characters in the movie (Billy Beane and Paul DePodesta) did not want to work with Hollywood.

- The filmmaker interviewed Ari instead and created the Paul character based on Ari’s real-life story.

- Ari happens to work at Aginity so we have a real ”Moneyball” for this project.

A Proof-of-Concept Demo for IBM Think Conference Talk

IBM Data Science @IBMDataScience — Following

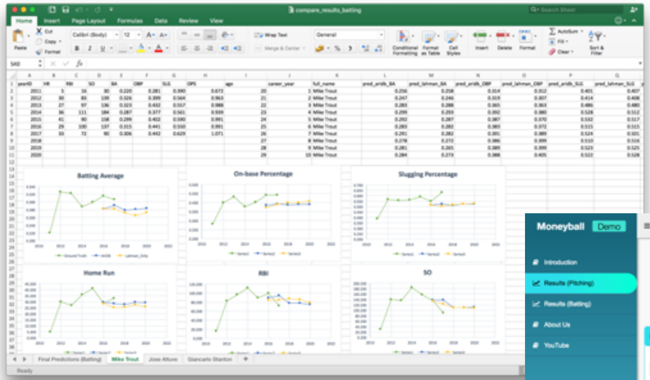
@DaithiOCiaran @arikaplan1 & @matlabulous are smoothly passing the mic back & forth to talk about their joint Moneyball project to a jam-packed room. #think2018 #machinelearning @Aginity @h2oai #DSX

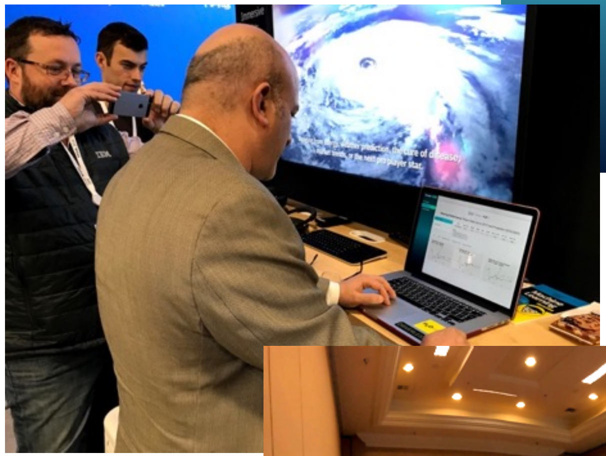
9:04 PM - 22 Mar 2018

H2O.ai

From PoC Demo to Real Moneyball

- **March 19** – AutoML Predictions finalized. Initial presentation in Excel.
- **March 20** – Version 1 of Shiny app. Ari used to app to validate some players he had in mind and recommended one player to his team.
- **March 21** – Multimillion-dollar contract finalized.
- **March 22** – Moneyball presentation at IBM Think





Framing the Business Problem for Machine Learning

Code on GitHub (without Ari's proprietary data)

https://github.com/woobe/moneyball

Search or jump to... /

Pull requests Issues Marketplace Explore

woobe / **moneyball**

Unwatch 2 | Star 1 | Fork 0

Code | Issues 0 | Pull requests 0 | Projects 0 | Wiki | Insights | Settings

Moneyball Demo (Public Version)

Edit

Add topics

7 commits | 1 branch | 0 releases | 1 contributor | Apache-2.0

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

woobe Added descriptions — Latest commit d630812 2 days ago

cache_data	Raw data from Lahman database	3 days ago
.gitignore	Initial commit	22 days ago
LICENSE	Initial commit	22 days ago
README.md	Added descriptions	2 days ago
step_1_data_munging.R	Data munging for Lahman data only	3 days ago
step_2_model_pitching.R	H2O AutoML Model Building Scripts	2 days ago
step_3_model_batting.R	H2O AutoML Model Building Scripts	2 days ago

README.md

Lahman Data

Player's information

birthYear	birthMonth	birthDay	birthCountry	birthState	birthCity
1991	8	7	USA	NJ	Vineland

nameFirst	nameLast	nameGiven	weight	height	bats	throws	debut	finalGame	retroID	bbrefID
Mike	Trout	Michael Nelson	235	74	R	R	2011-07-08	2017-10-01	troum001	troutmi01

Player's past performance (batting in this case)

d_bat_raw

Filter

	playerID	yearID	stint	teamID	lgID	G	AB	R	H	2B	3B	HR	RBI	SB	CS	BB	SO	IBB	HBP	SH	SF	GIDP
95484	troutmi01	2011	1	LAA	AL	40	123	20	27	6	0	5	16	4	0	9	30	0	2	0	1	2
96904	troutmi01	2012	1	LAA	AL	139	559	129	182	27	8	30	83	49	5	67	139	4	6	0	7	7
98308	troutmi01	2013	1	LAA	AL	157	589	109	190	39	9	27	97	33	7	110	136	10	9	0	8	8
99744	troutmi01	2014	1	LAA	AL	157	602	115	173	39	9	36	111	16	2	83	184	6	10	0	10	6
101226	troutmi01	2015	1	LAA	AL	159	575	104	172	32	6	41	90	11	7	92	158	14	10	0	5	11
102712	troutmi01	2016	1	LAA	AL	159	549	123	173	32	5	29	100	30	7	116	137	12	11	0	5	5
104195	troutmi01	2017	1	LAA	AL	114	402	92	123	25	3	33	72	22	4	94	90	15	7	0	4	8

Lahman Data Framed as a ML problem

Player Attributes

yearID	teamID	lgID	weight	height	bats	throws	birthYear	birthCountry	birthState	birthCity	age	career_year
2011	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	20	1
2012	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	21	2
2013	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	22	3
2014	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	23	4
2015	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	24	5
2016	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	25	6
2017	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	26	7
2018	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	27	8
2019	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	28	9
2020	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland	29	10

Past Performance Sliding Windows + Other Stats

last1_HR	last2_HR	last3_HR	last4_HR	last5_HR	avg_last2_HR	avg_last3_HR	avg_last4_HR	avg_last5_HR
NA	NA	NA	NA	NA	NaN	NaN	NaN	NaN
5	NA	NA	NA	NA	5.0	5.00000	5.00000	5.00000
30	5	NA	NA	NA	17.5	17.50000	17.50000	17.50000
27	30	5	NA	NA	28.5	20.66667	20.66667	20.66667
36	27	30	5	NA	31.5	31.00000	24.50000	24.50000
41	36	27	30	5	38.5	34.66667	33.50000	27.80000
29	41	36	27	30	35.0	35.33333	33.25000	32.60000
33	29	41	36	27	31.0	34.33333	34.75000	33.20000
33	33	29	41	36	33.0	31.66667	34.00000	34.40000
33	33	33	29	41	33.0	33.00000	32.00000	33.80000

One of the Targets

yearID	HR
2011	5
2012	30
2013	27
2014	36
2015	41
2016	29
2017	33
2018	NA
2019	NA
2020	NA

Training

Validation

Forecast

No data. Used 2017 value. Not perfect (a quick hack).

Approach One: Learning from **Lahman** only

H_2O AutoML:
Learn the Pattern

Lahman: Age, Height, Weight ...
Historical Performance Stats

$$\begin{bmatrix} \quad \mathbf{X} \quad \end{bmatrix} \begin{bmatrix} \mathbf{y} \end{bmatrix} \xrightarrow{f} \begin{bmatrix} \hat{\mathbf{y}} \end{bmatrix}$$

Sliding Windows (Stats from previous n years)
About **300 Lahman** Features

Home Runs
Batting Average
...

Predictions

Approach Two: Learning from **Lahman** & **AriDB**

H_2O AutoML:
Learn the Pattern

Lahman: Age, Height, Weight …
Historical Performance Stats

$$\left[\mathbf{X}\right] \quad \left[\mathbf{y}\right] \xrightarrow{f} \left[\hat{\mathbf{y}}\right]$$

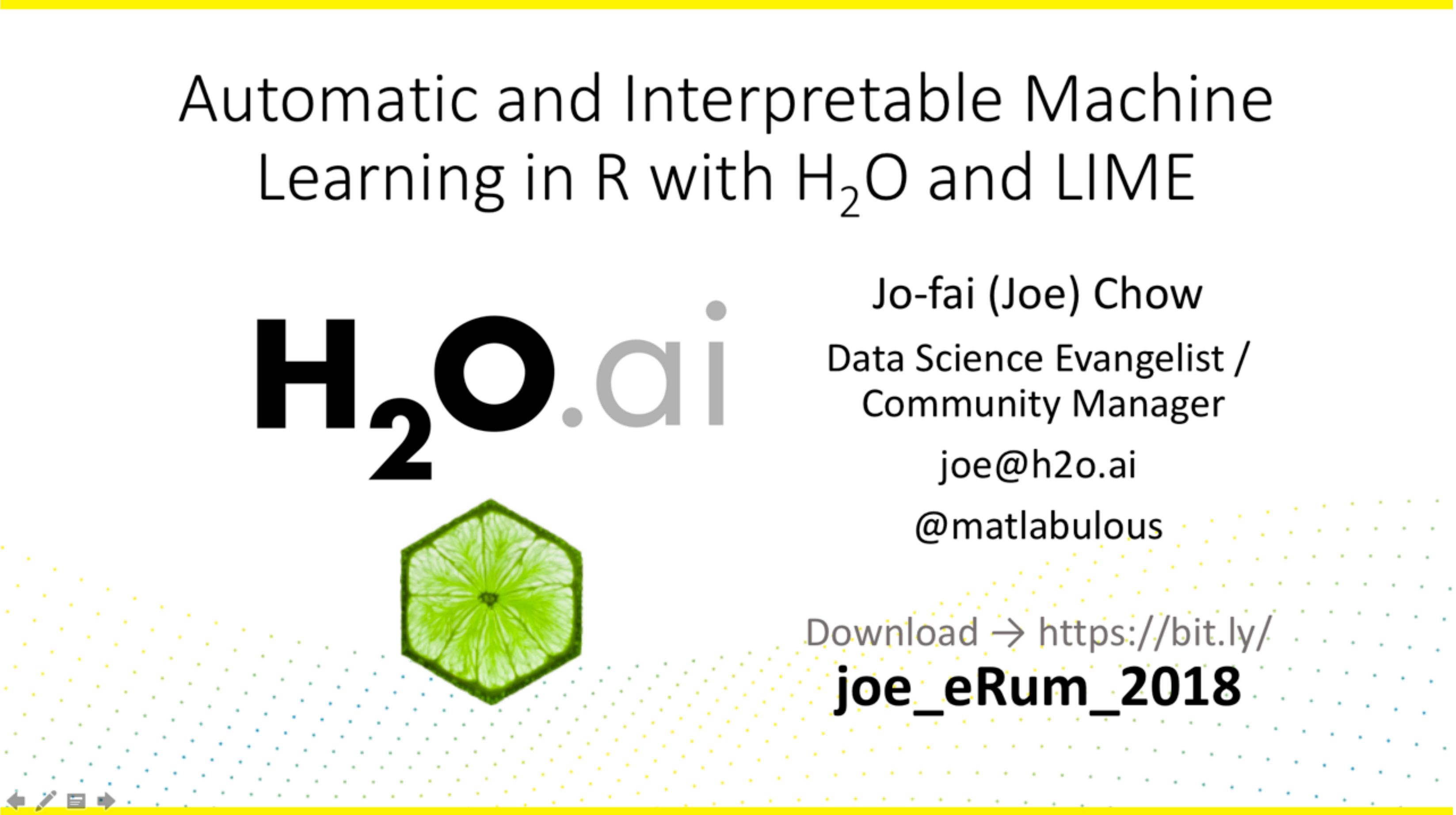
AriDB: Fastball, curveball,
slider, velocity …

Sliding Windows (Stats from previous n years)
About **300 Lahman** Features +
200 AriDB Features

Home Runs
Batting Average
…

Predictions

H_2O AutoML and LIME Materials for eRum Workshop



LIME

Reference: https://github.com/thomasp85/lime

```r
# Install 'lime' from CRAN
install.packages('lime')
```

H2O

Reference: https://www.h2o.ai/download/

```r
# Install 'h2o' from CRAN
install.packages('h2o')
```

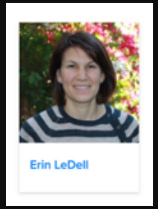
… and **mlbench** for datasets

About H_2O AutoML

Automatic Machine Learning with H_2O

http://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html

H_2O AutoML



Erin LeDell

Navdeep Gill

H_2O Team

H_2O-3 Algorithms Overview

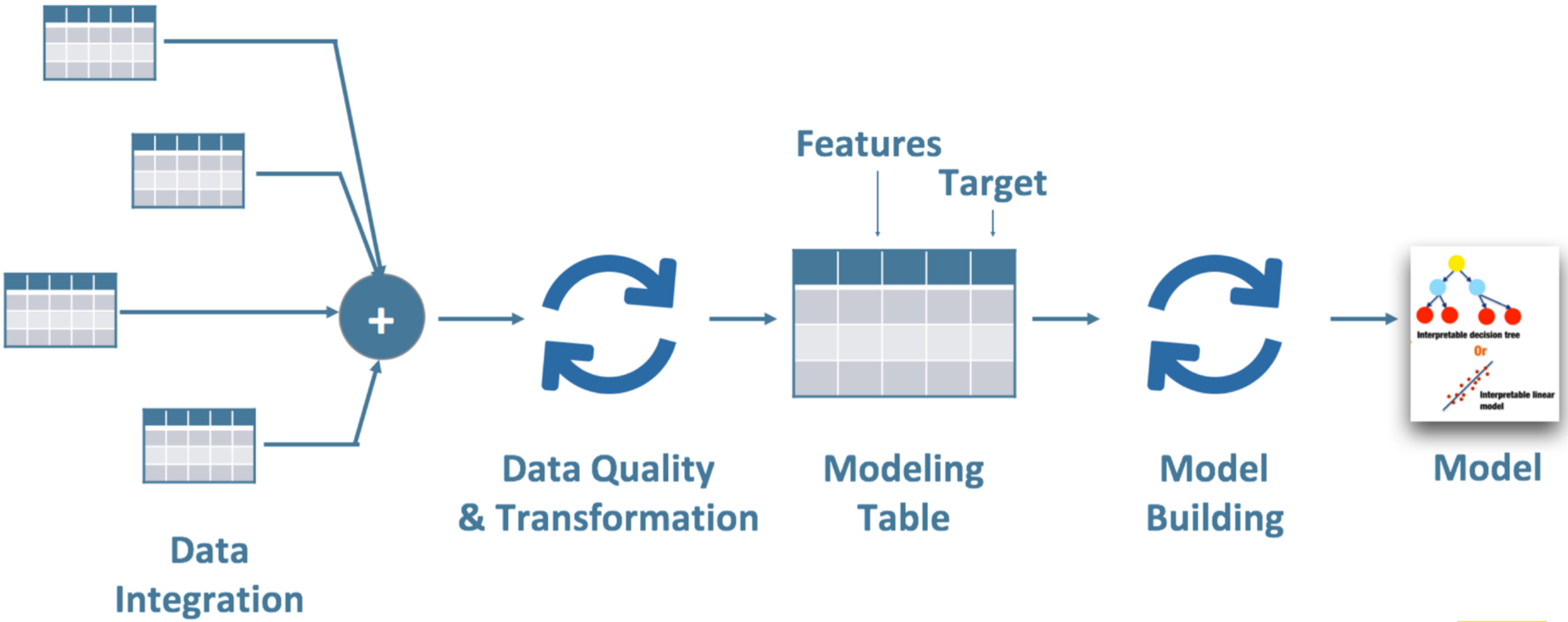
Supervised Learning

Statistical Analysis	• **Generalized Linear Models**: Binomial, Gaussian, Gamma, Poisson and Tweedie • **Naïve Bayes**
Ensembles	• **Distributed Random Forest**: Classification or regression models • **Gradient Boosting Machine**: Produces an ensemble of decision trees with increasing refined approximations
Deep Neural Networks	• **Deep learning**: Create multi-layer feed forward neural networks starting with an input layer followed by multiple layers of nonlinear transformations

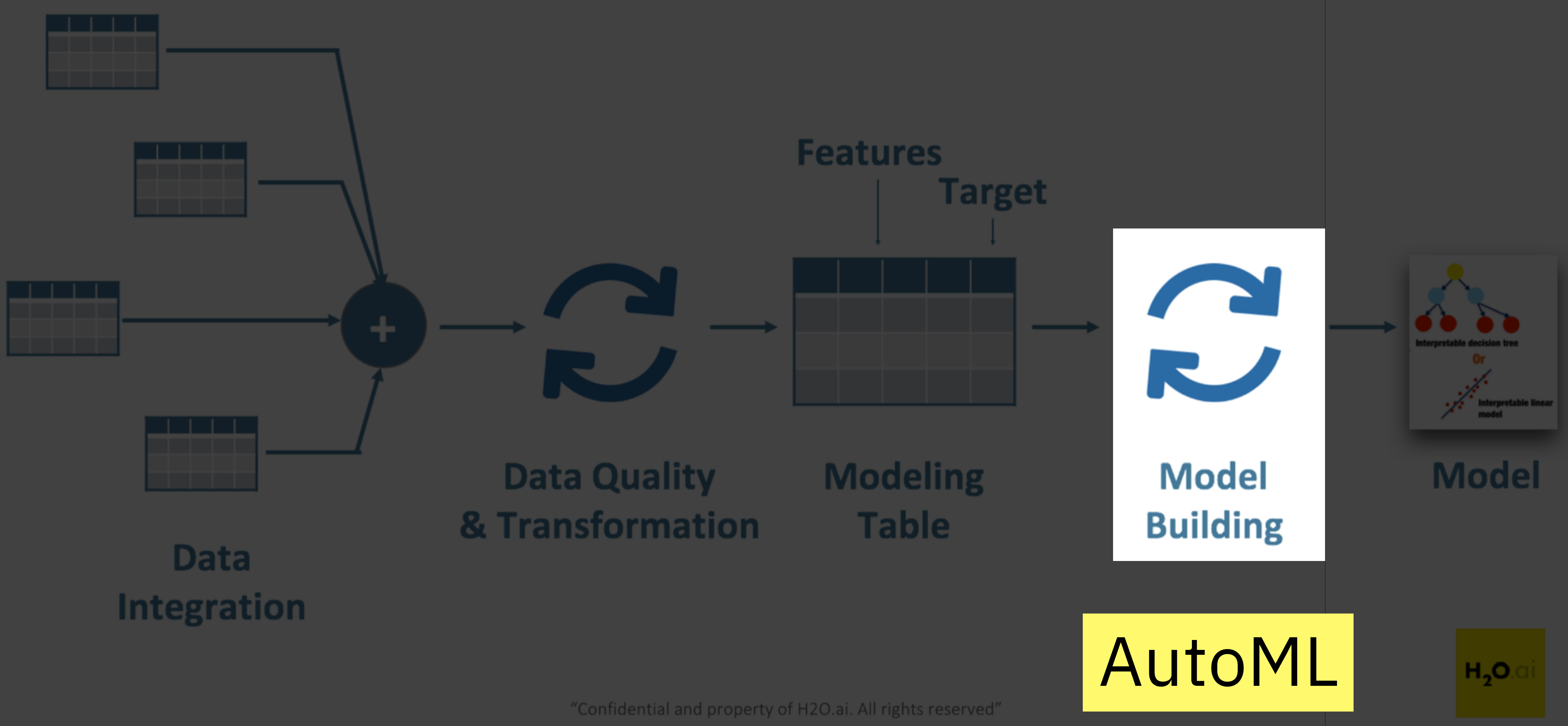
Unsupervised Learning

Clustering	• **K-means**: Partitions observations into k clusters/groups of the same spatial size. Automatically detect optimal k
Dimensionality Reduction	• **Principal Component Analysis**: Linearly transforms correlated variables to independent components • **Generalized Low Rank Models:** extend the idea of PCA to handle arbitrary data consisting of numerical, Boolean, categorical, and missing data
Anomaly Detection	• **Autoencoders**: Find outliers using a nonlinear dimensionality reduction using deep learning

Typical Enterprise Machine Learning Workflow



Typical Enterprise Machine Learning Workflow



AutoML Code

```
# H2O AutoML with Lahman only
automl_lahman = h2o.automl(x = features,
                           y = targets[n_target],
                           training_frame = h_train,
                           validation_frame = h_valid,
                           max_models = 10, # increase this to allow more models
                           max_runtime_secs = 120, # increase this to allow more time
                           stopping_metric = "RMSE",
                           stopping_rounds = 3,
                           seed = n_seed,
                           exclude_algos = c("DeepLearning"), # you can exclude any algo
                           project_name = paste0("AutoML_Lahman", targets[n_target]))
```

AutoML Results

```
H2ORegressionMetrics: stackedensemble
** Reported on cross-validation data. **
** 5-fold cross-validation on training data (Metrics computed for combined holdout predictions) **

MSE:  0.00246453
RMSE:  0.04964404
MAE:  0.03335875
RMSLE:  0.04124294
Mean Residual Deviance :  0.00246453



Slot "leaderboard":
                                             model_id mean_residual_deviance     rmse      mae    rmsle
1 StackedEnsemble_BestOfFamily_0_AutoML_20180615_040834               0.002465 0.049644 0.033359 0.041243
2    StackedEnsemble_AllModels_0_AutoML_20180615_040834               0.002467 0.049669 0.033367 0.041265
3             GLM_grid_0_AutoML_20180615_040834_model_0               0.002480 0.049802 0.033560 0.041401
4             GBM_grid_0_AutoML_20180615_040834_model_4               0.002486 0.049856 0.033707 0.041373
5             GBM_grid_0_AutoML_20180615_040834_model_2               0.002564 0.050638 0.034346 0.042008
6             GBM_grid_0_AutoML_20180615_040834_model_1               0.002569 0.050684 0.034261 0.042022

[12 rows x 5 columns]
```

About Machine Learning Interpretability

LIME (Local Interpretable Model-Agnostic Explanations)

… and more

Acknowledgement

- **Marco Tulio Ribeiro**: Original LIME Framework and Python package

- **Thomas Lin Pedersen**: LIME R package
- **Matt Dancho**: LIME + H2O AutoML example + LIME R package improvement

- **Kasia Kulma**: LIME + H2O AutoML example

Why Should I Trust Your Model?

System that performs behaviour but you don’t know how it works

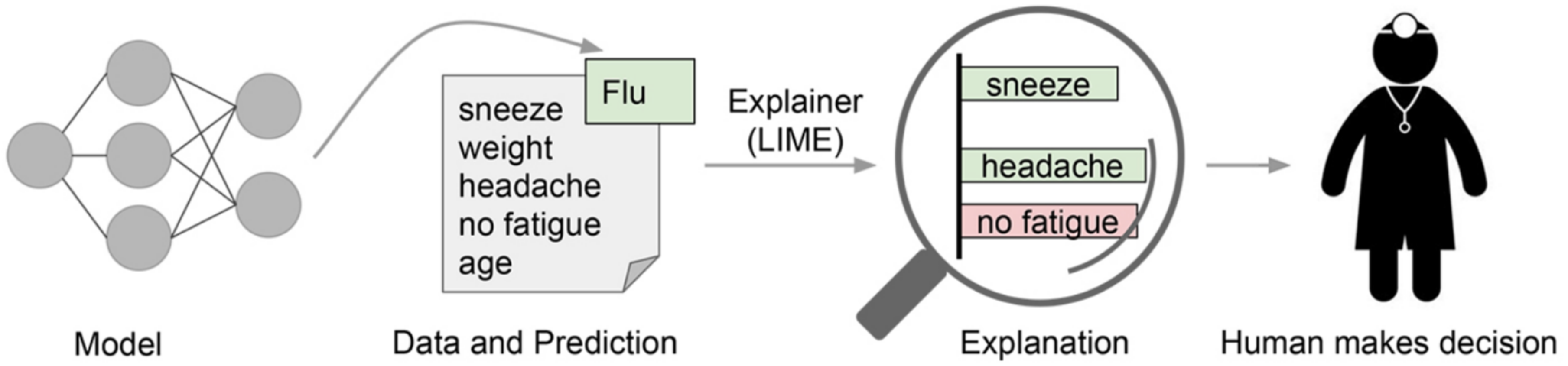


Figure 1. Explaining individual predictions to a human decision-maker. Source: Marco Tulio Ribeiro.

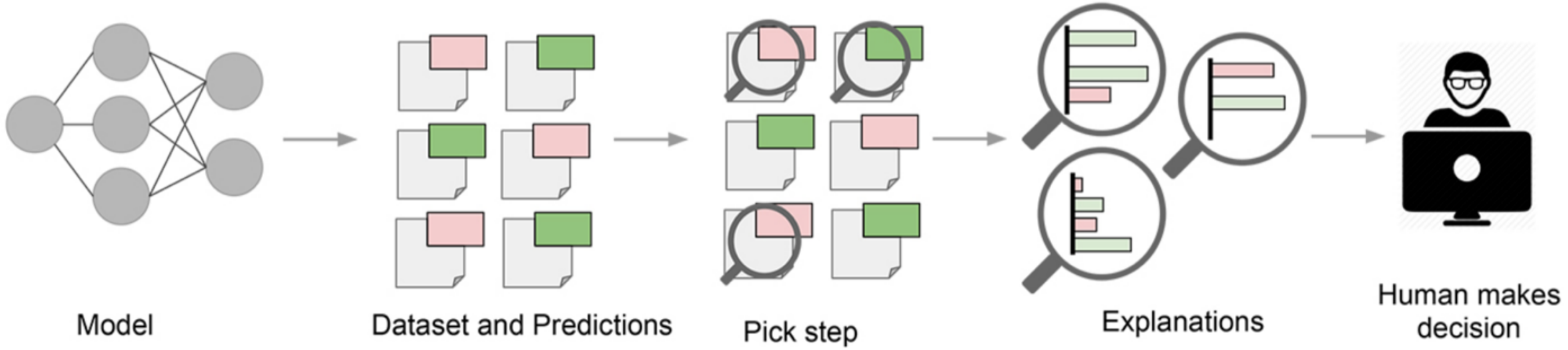
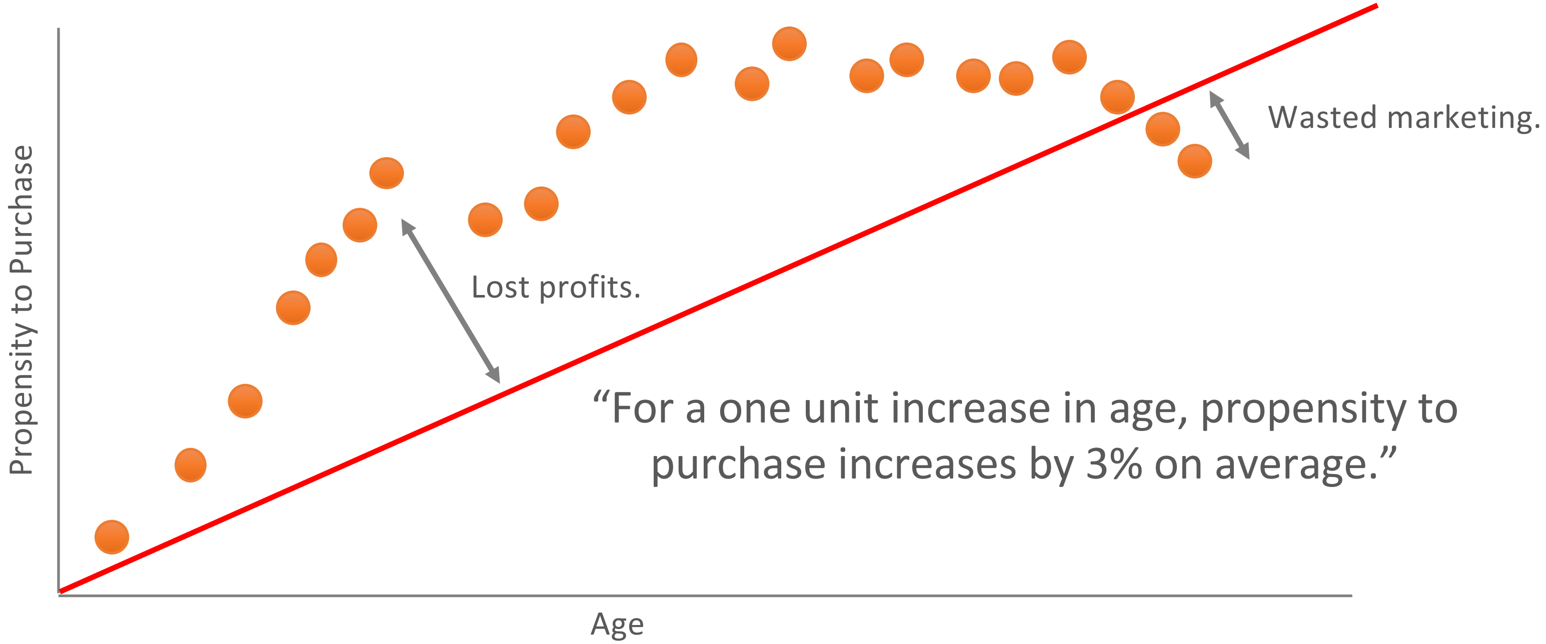


Figure 2. Explaining a model to a human decision-maker. Source: Marco Tulio Ribeiro.

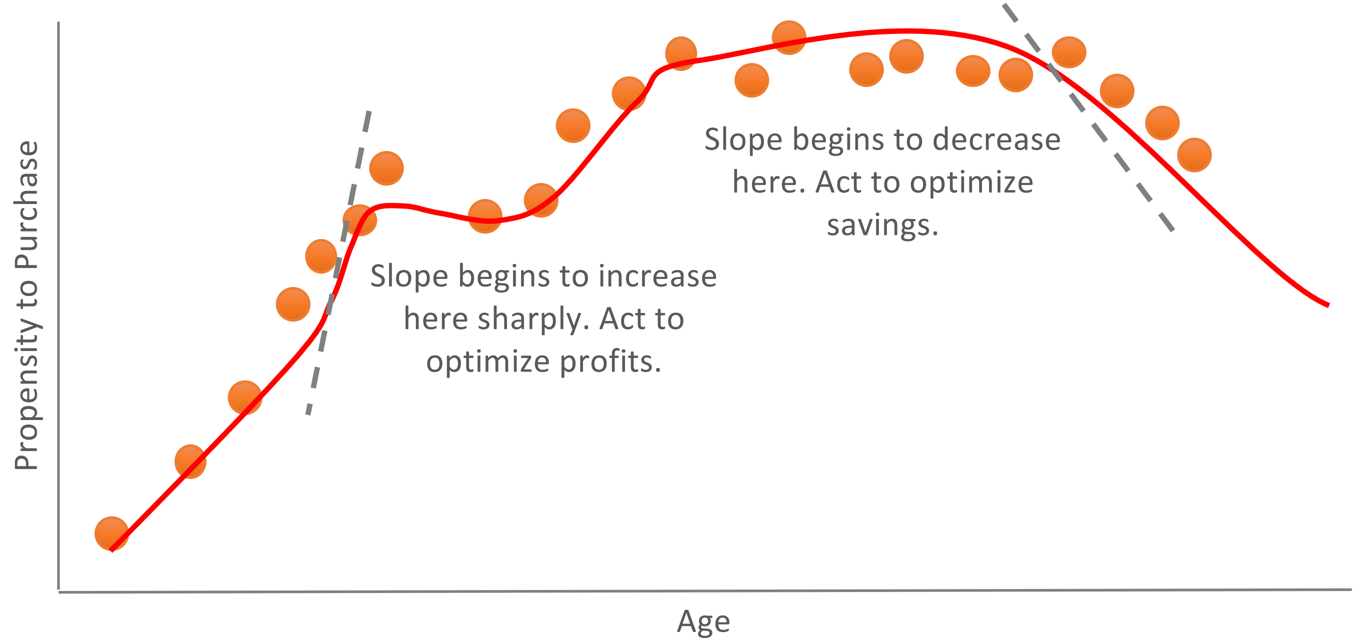
Linear Models

Exact explanations for approximate models.



Machine Learning

Approximate explanations for exact models.



H2O.ai

Local Interpretable Model-Agnostic Explanations LIME - How does it work?

Theory

- LIME approximates model locally as logistic or linear model
- Repeats process many times
- Outputs features that are most important to local models

Outcome

- Approximate reasoning
- Complex models can be interpreted
  - Neural nets, Random Forest, Ensembles etc.

From my eRum workshop – see bit.ly/joe_eRum_2018

4 Explain the Model

4.1 Step 1: Create an `explainer`

```
explainer = lime::lime(x = as.data.frame(h_train[, features]),
                       model = model_automl@leader)
```

4.2 Step 2: Turn `explainer` into `explanations`

```
# Extract one sample (change `1` to any row you want)
d_samp = as.data.frame(h_test[1, features])
```

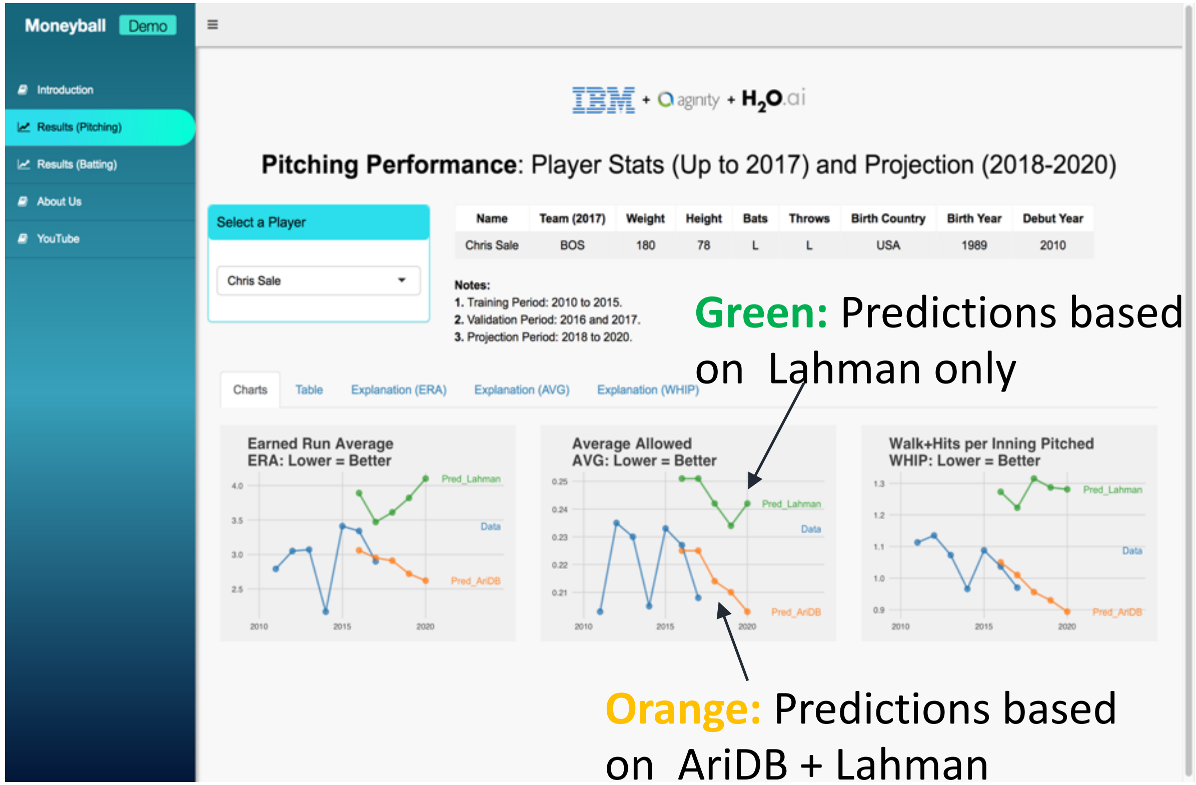
```
# Assign a specifc row name (for better visualization)
row.names(d_samp) = "Sample 1"
```

```
# Create explanations
explanations = lime::explain(x = d_samp,
                             explainer = explainer,
                             n_permutations = 5000,
                             feature_select = "auto",
                             n_features = 13) # Look top x features
```

4.3 Look at Explanations (Bar Chart)

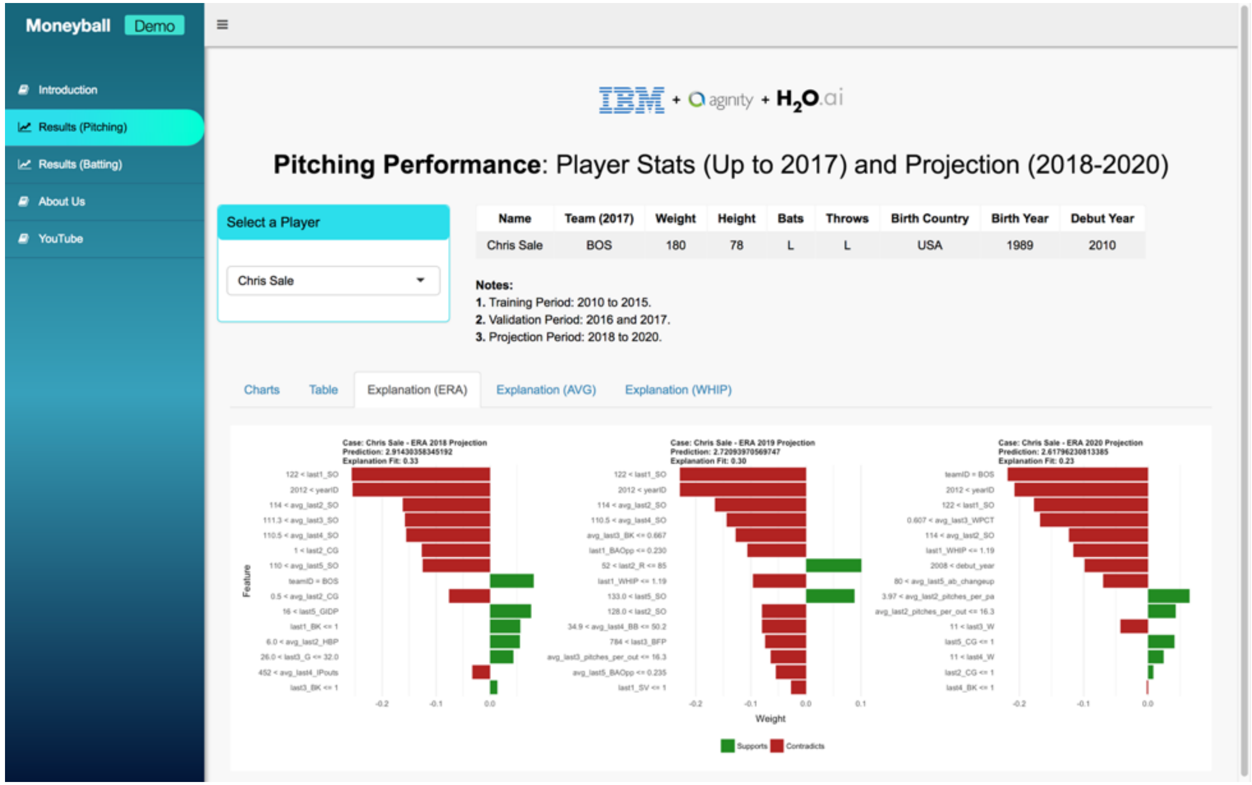
```
lime::plot_features(explanations, ncol = 1)
```

Live Demo (again)



Green: Predictions based on Lahman only

Orange: Predictions based on AriDB + Lahman



H2O.ai

Other H_2O News

H_2O Products

H2O.ai

In-Memory, Distributed Machine Learning Algorithms with H2O Flow GUI

Spark + H_2O SPARKLING WATER

H2O AI Open Source Engine Integration with Spark

H2O4GPU

Lightning Fast machine learning on GPUs

DRIVERLESS AI

Automatic feature engineering, machine learning and interpretability

Steam

Secure multi-tenant H2O clusters

H2O.ai

Algorithms on H2O-3 (CPU)

Supervised Learning

Statistical Analysis	• Generalized Linear Models: Binomial, Gaussian, Gamma, Poisson and Tweedie • Naïve Bayes
Ensembles	• Distributed Random Forest: Classification or regression models • Gradient Boosting Machine: Produces an ensemble of decision trees with increasing refined approximations
Deep Neural Networks	• Deep learning: Create multi-layer feed forward neural networks starting with an input layer followed by multiple layers of nonlinear transformations

Unsupervised Learning

Clustering	• K-means: Partitions observations into k clusters/groups of the same spatial size. Automatically detect optimal k
Dimensionality Reduction	• Principal Component Analysis: Linearly transforms correlated variables to independent components • Generalized Low Rank Models: extend the idea of PCA to handle arbitrary data consisting of numerical, Boolean, categorical, and missing data
Anomaly Detection	• Autoencoders: Find outliers using a nonlinear dimensionality reduction using deep learning

Algorithms on H_2O4GPU (more to come)

Supervised Learning

Statistical Analysis

- Generalized Linear Models: Binomial, Gaussian, Gamma, Poisson and Tweedie
- Naïve Bayes

Ensembles

- Distributed Random Forest: Classification or regression models
- Gradient Boosting Machine: Produces an ensemble of decision trees with increasing refined approximations

Deep Neural Networks

- Deep learning: Create multi-layer feed forward neural networks starting with an input layer followed by multiple layers of nonlinear transformations

Unsupervised Learning

Clustering

- K-means: Partitions observations into k clusters/groups of the same spatial size. Automatically detect optimal k

Dimensionality Reduction

- Principal Component Analysis: Linearly transforms correlated variables to independent components
- Generalized Low Rank Models: extend the idea of PCA to handle arbitrary data consisting of numerical, Boolean, categorical, and missing data

Anomaly Detection

- Autoencoders: Find outliers using a nonlinear dimensionality reduction using deep learning

H2O4GPU now available in R

BY **ERIN LEDELL** ON MARCH 27, 2018 — **0 COMMENTS**

In September, H2O.ai released a new open source software project for GPU machine learning called H2O4GPU. The initial release (blog post here) included a Python module with a scikit-learn compatible API, which allows it to be used as a drop-in replacement for scikit-learn with support for GPUs on selected (and ever-growing) algorithms. We are proud to announce that the same collection of GPU algorithms is now available in R, and the **h2o4gpu** R package is available on CRAN.

H2O4GPU

https://github.com/h2oai/h2o4gpu

From Kaggle Grand Masters' Recipes to Production Ready in a Few Clicks

BY JO-FAI CHOW ON MAY 9, 2018 — 0 COMMENTS — EDIT

Introducing Accelerated Automatic Pipelines in H2O Driverless AI

At H2O, we work really hard to make machine learning fast, accurate, and accessible to everyone. With H2O Driverless AI, users can leverage years of world-class, Kaggle Grand Masters experience and our GPU-accelerated algorithms (H2O4GPU) to produce top quality predictive models in a fully automatic and timely fashion.

In our most recent release (version 1.1), we are going one step further to streamline the deployment process with MOJO (**M**odel **O**b**J**ect, **O**ptimized). Inherited from our popular H2O-3 platform, MOJO is a highly optimized, low-latency scoring engine that is easily embeddable in any Java environment. With automatic pipeline generation in Driverless AI, users can go from automatic machine learning to production ready in just a few clicks. This blog post illustrates the usage of MOJO in Driverless AI with a simple example.

blog.h2o.ai

Easing the Pain Points in a Machine Learning Workflow

In a typical enterprise machine learning workflow, there are many things that could go wrong due to human errors, bad data science practices, different tools/infrastructure, incompatible code, lack of testing, versioning, communication and so on.

meetup

Start a new group | Explore | Messages | Notifications

19 JUN

Tuesday, June 19, 2018

June #LondonAI: XAI, Neural Style Transfer & e-Learning (External Reg Required)

Hosted by Jo-fai Chow
From London Artificial Intelligence & Deep Learning
Public group

You're going

Share:

Details

We have a new venue for the meetup!

I love this community! Not long after I sent out the group email about the emergency change of venue, many of you reached out to me and offered help. I really appreciate this :)

Thanks to Michael James from Yoox-Net-A-Porter, we will host our meetup at their Tech Hub in White City.

Please register with your full name and email for security purposes. Here is the Eventbrite link:
https://www.eventbrite.com/e/june-londonai-meetup-tickets-46242661044

Agenda:

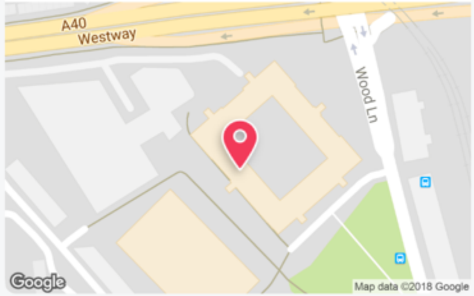
- 6:00 to 6:30pm Pizza and Drinks (please come early if possible)
- Introduction by H2O.ai (5 mins)
- Introduction by Yoox-Net-A-Porter Cognitive Commerce Team (15 mins)

Tech Talks:
- Explainable Artificial Intelligence (XAI) by Torgyn Shaikhina (20 mins)
- Neural Style Transfer by Ambroise Laurent (20 mins)
- Applying AI/ML to e-Learning by Shabbir Mookhtiar (20 mins)

Organizer tools

Tuesday, June 19, 2018
6:00 PM to 9:00 PM
Add to calendar

YOOX NET A PORTER GROUP
Tech Hub, 2nd Floor, Building 6 (The Mediaworks Building), Wood Lane, White City Place, London, W127TU · London



London Meetup
June 19 (Next Tuesday)

Danke!

- Organizers & Sponsors

- Code, Slides & Documents
  - bit.ly/h2o_meetups
  - bit.ly/joe_eRum_2018
  - docs.h2o.ai
- Contact
  - joe@h2o.ai
  - @matlabulous
  - github.com/woobe
- Please search/ask questions on **Stack Overflow**
  - Use the tag `**h2o**` (not h2 zero)