



# Will You Buy More?

## Data Science Course Project



JESSICA • GROWTH ANALYST • ATLASSIAN



DEALS STARTING AT 30% OFF | 30% OFF TABLETOP | 25% OFF CANDLES + HOME FRAGRANCES | UP TO 70% OFF MARKDOWNS

[Home](#) > [Furniture](#) > [Nightstands + Dressers + Wardrobes](#) > [Dressers](#) > Mid-Century 6-Drawer Dresser - Acorn



## Mid-Century 6-Drawer Dresser - Acorn

•

**\$999**

White Glove Service + \$100

- Home Delivery

1



[ADD TO REGISTRY](#)

[ADD TO CART](#)

[ADD TO WISHLIST](#)

Recently





Search Facebook



Jessica

Home

 Jessica Huffman Edit Profile

## FAVORITES

 News Feed Messages Events

5

 Saved

4

 You Can Run - Sea... Family Shops Buy and Sell Groups

## GROUPS

 Impala Racing Team

2

 Brunch Biddies

7

 Donner Party Swim

20+

 Discover Groups Create Group

## FRIENDS

 Recreation

## APPS

 Live Video

What's on your mind?

 Photo / Video Photo Album

...

 Friends ▾

Post

**Sarah Glushko Karp** shared Colin Ebeling's photo.7 hrs · 

Ballot sent from Oz. Go get 'em. #imwithher

 19 5 event invites Jorge Moreno and 5 others

## TRENDING

 Central Catholic High School

44K people talking about this

 Angelina Jolie

170K people talking about this

 Condoleezza Rice

62K people talking about this

 See More

## SPONSORED

Create Ad

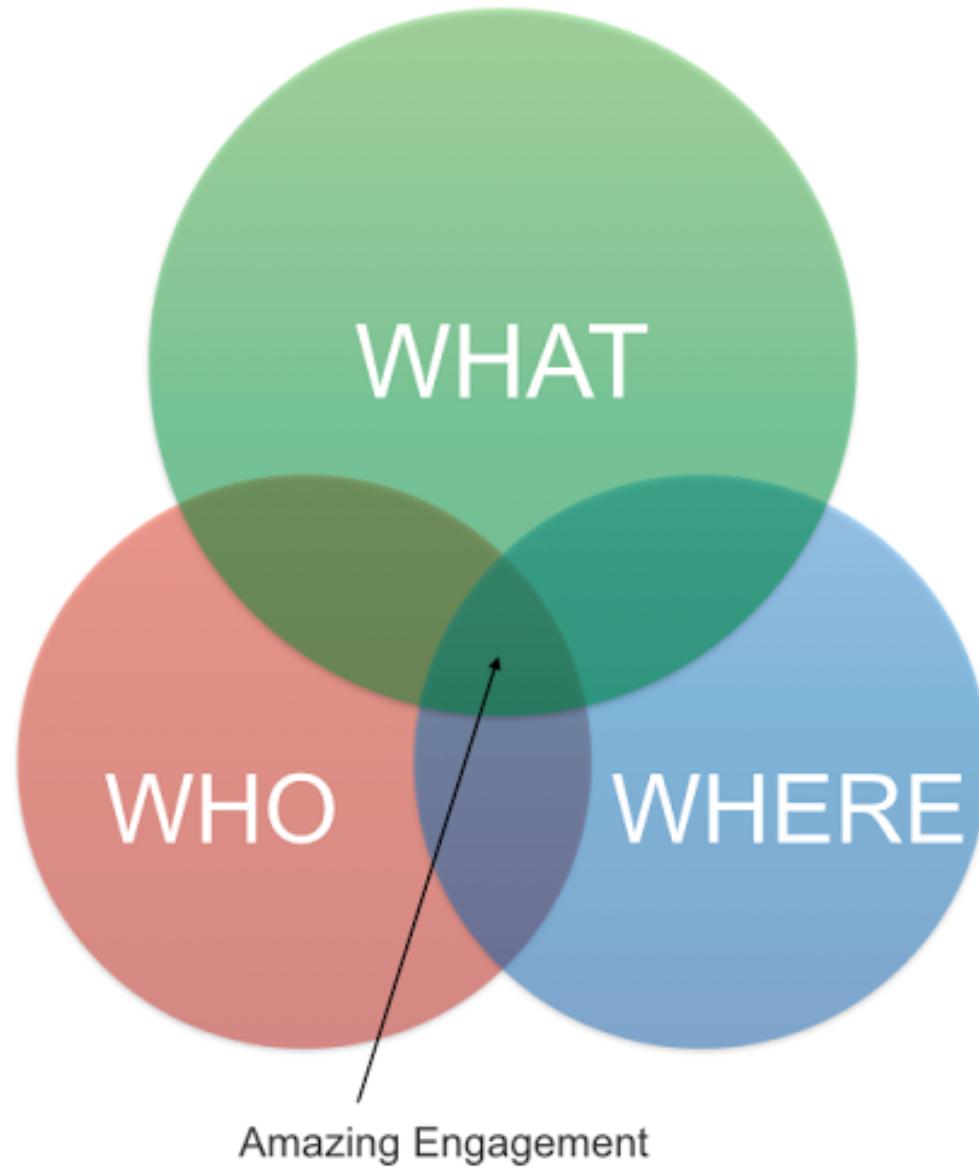
**Shop West Elm**

westelm.com

Shop Mid-Century 6-Drawer Dresser - Acorn and more at West Elm.

# Personalized Contextual Messaging







 JIRA Software +  Bitbucket

# Goal

Can we predict if a **new JIRA instance will purchase Bitbucket** based on **JIRA in-product behavior and license characteristics?**



Null Hypothesis

11%

of new JIRA instances  
purchase Bitbucket



# Data

- Customer purchased JIRA between 2014-2016
  - JIRA Cloud licenses (n=16,560)
  - Bitbucket Cloud licenses
- In Product Analytics (user behavior)
- NPS Survey Data
- Segment Data (website analytics)

# Model Overview

- **Type:** Supervised Classification
- **Response:**
  - Y= Co\_ownership
- **Features:**
  - Normalized MAU (percent of users using license tier, MAU at month 3)
  - Github integration
  - Bundled
  - Owning other Atlassian Products (Confluence, Sourcetree, Bamboo, Service Desk)
  - JIRA Evaluation Score
  - In product user behavior
    - comments, mentions, issues, projects, sprints, workflows, release hub, versions, epics, scrum, kanban

Not yet included: months owning JIRA, JIRA Highly Active MAU in month 3, JIRA NPS, marketing channel, evaluation counts

```
##Observations
```

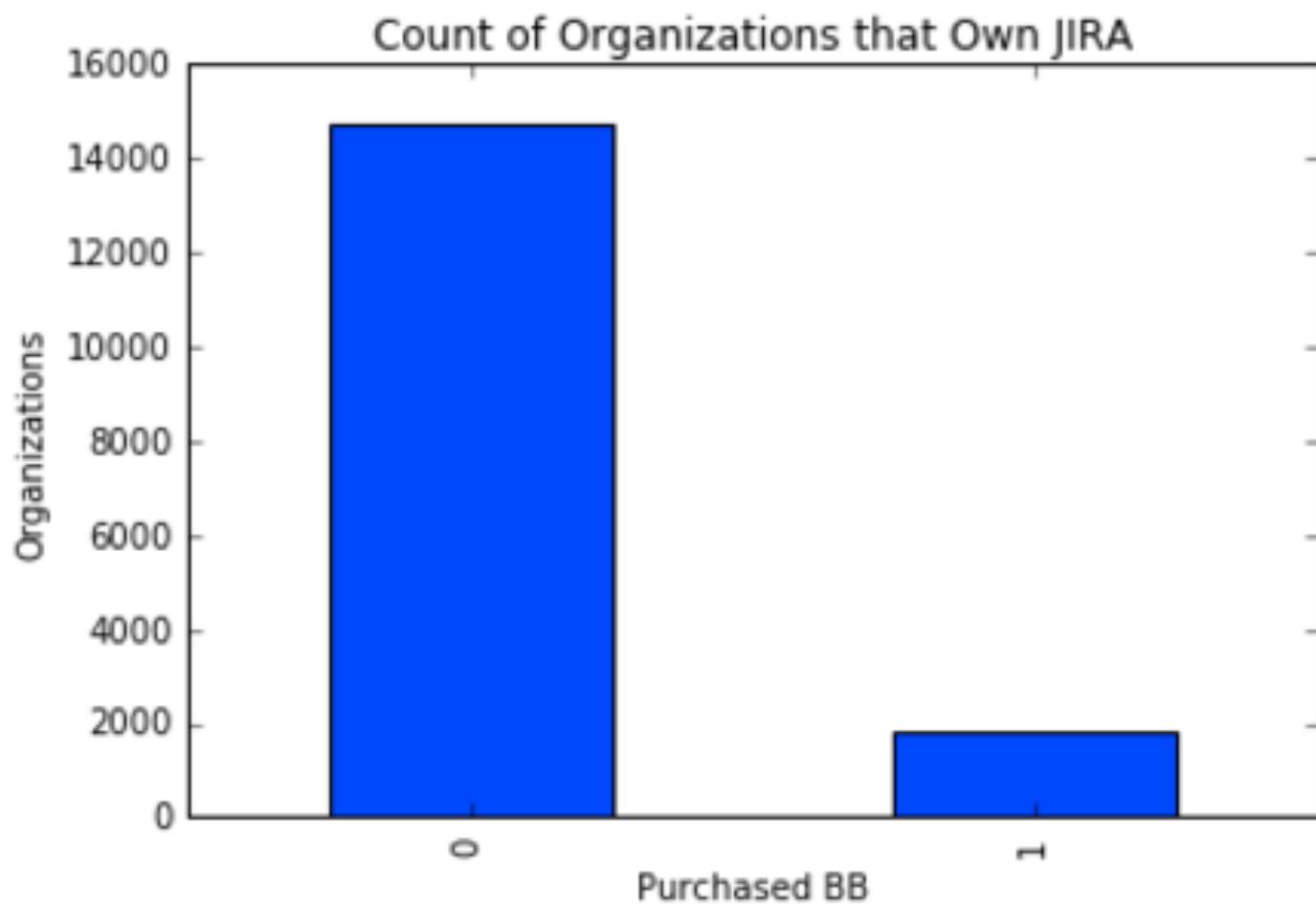
```
license_data = full_data[['jira_start_date','jira','jira_user_limit','jira_mau','normalized_mau','GH_integrated','eval_score','bundled','co_ownership']]  
license_data
```

	jira_start_date	jira	jira_user_limit	jira_mau	normalized_mau	GH_integrated	eval_score	bundled	co_ownership
0	2014-10-27	JIRA Software (Cloud)	10	3	0.300000	0	0.052945	0	0
1	2015-11-04	JIRA Software (Cloud)	10	8	0.800000	0	0.076112	1	1
2	2015-03-04	JIRA Software (Cloud)	10	0	0.000000	0	0.076053	1	0
3	2014-05-01	JIRA Software (Cloud)	50	16	0.320000	0	0.118175	1	1
4	2014-08-31	JIRA Software (Cloud)	10	0	0.000000	0	0.009802	0	0
5	2015-05-11	JIRA Software (Cloud)	10	4	0.400000	0	0.022192	0	0
6	2015-11-02	JIRA Software (Cloud)	10	4	0.400000	0	0.137785	1	0
7	2014-03-12	JIRA Software (Cloud)	50	40	0.800000	1	0.152956	0	0

```
##Observations
```

```
behavior_data = full_data[['jira_start_date','normalized_mau','prod_comment','prod_mention','prod_release','prod_epic','behavior_data']]
```

	jira_start_date	normalized_mau	prod_comment	prod_mention	prod_release	prod_epics	prod_kanban	prod_scrum	co_ownership
0	2014-10-27	0.300000	0	0	0	0	0	0	0
1	2015-11-04	0.800000	1	0	1	0	1	1	1
2	2015-03-04	0.000000	0	0	0	0	0	0	0
3	2014-05-01	0.320000	0	0	0	0	0	0	1
4	2014-08-31	0.000000	0	0	0	0	0	0	0
5	2015-05-11	0.400000	0	0	0	0	0	0	0
6	2015-11-02	0.400000	1	0	0	1	0	1	0



# Correlation among Predictors

	co_ownership	GH_integrated	normalized_mau	bundled	sourcetree	bamboo
co_ownership	1.000000	-0.076046	0.135003	0.141695	0.286117	0.199331
GH_integrated	-0.076046	1.000000	0.111278	0.080881	0.042899	0.031682
normalized_mau	0.135003	0.111278	1.000000	0.083310	0.130565	0.055763
bundled	0.141695	0.080881	0.083310	1.000000	0.134065	0.124010
sourcetree	0.286117	0.042899	0.130565	0.134065	1.000000	0.175737
bamboo	0.199331	0.031682	0.055763	0.124010	0.175737	1.000000



# Correlation among Predictors

	co_ownership	prod_sprintcomplete	prod_workflow	prod_versionrelease	prod_epics	prod_kanban	prod_scrum
co_ownership	1.000000	-0.014735	-0.022886	-0.006078	-0.012976	-0.029599	-0.064528
prod_sprintcomplete	-0.014735	1.000000	0.568738	0.459017	0.616198	0.517685	0.710945
prod_workflow	-0.022886	0.568738	1.000000	0.505619	0.451815	0.615111	0.665780
prod_versionrelease	-0.006078	0.459017	0.505619	1.000000	0.388931	0.565644	0.479023
prod_epics	-0.012976	0.616198	0.451815	0.388931	1.000000	0.438494	0.527532
prod_kanban	-0.029599	0.517685	0.615111	0.565644	0.438494	1.000000	0.640396
prod_scrum	-0.064528	0.710945	0.665780	0.479023	0.527532	0.640396	1.000000



# Results

Null Hypothesis: 0.8896

	Type	CV	Accuracy	ROC/AUC
<b>Model 1</b>	KNN	10	0.8917	0.8027
<b>Model 2</b>	Log Reg	10	0.8921	0.8117
<b>Model 3</b>	NB	10	0.8854	0.7326
<b>Model 4</b>	RF	10	0.8675	0.6580
<b>Model 5</b>	Decision Tree	5	0.8742	0.7361



# Results

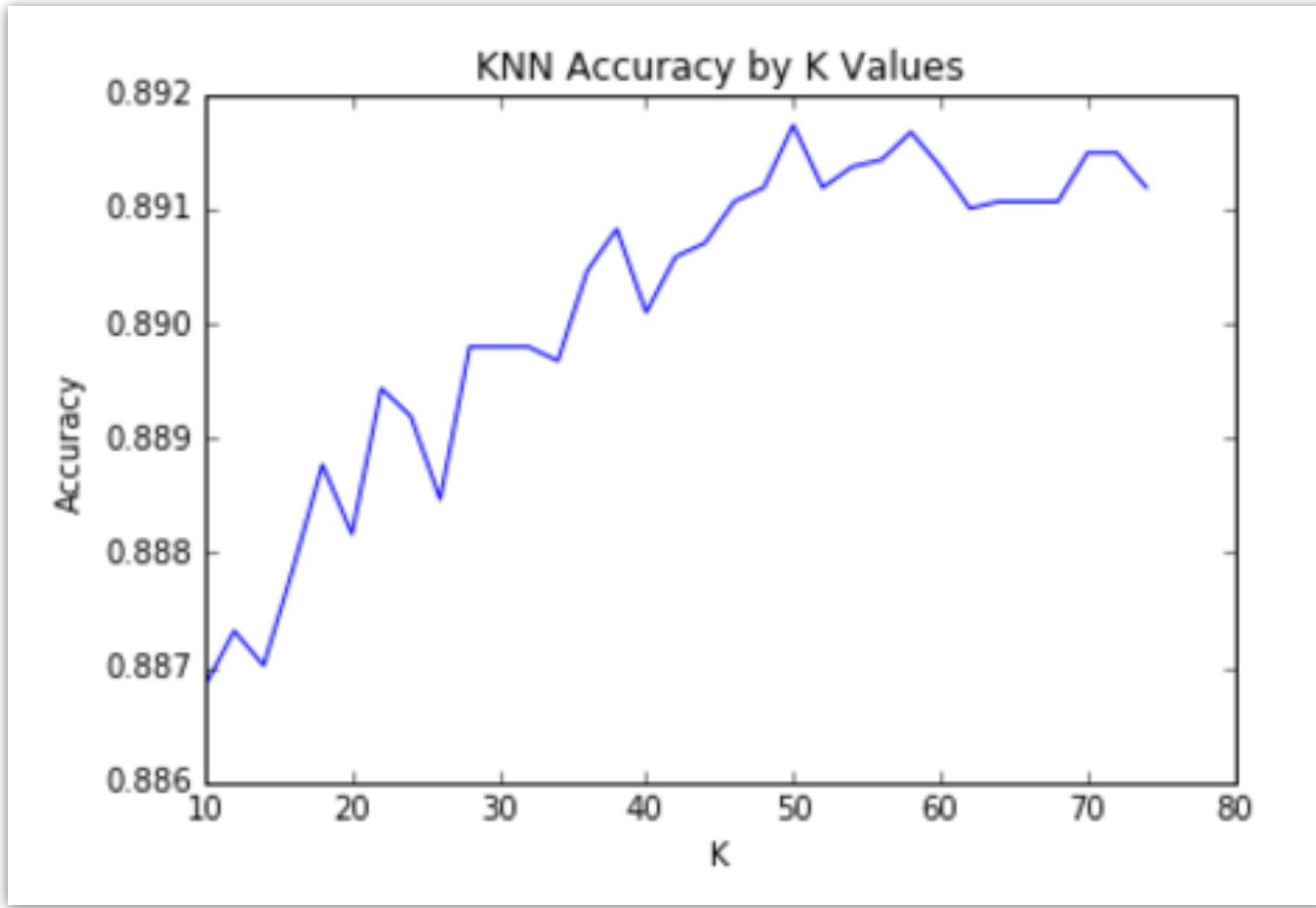
Null Hypothesis: 0.8896

	Type	CV	Accuracy	ROC/AUC
<b>Model 1</b>	KNN	10	<b>0.8922</b>	<b>0.8027</b>
<b>Model 2</b>	Log Reg	10	<b>0.8921</b>	<b>0.8117</b>
<b>Model 3</b>	NB	10	0.8854	0.7326
<b>Model 4</b>	RF	10	0.8675	0.6580
<b>Model 5</b>	Decision Tree	5	0.8742	0.7361

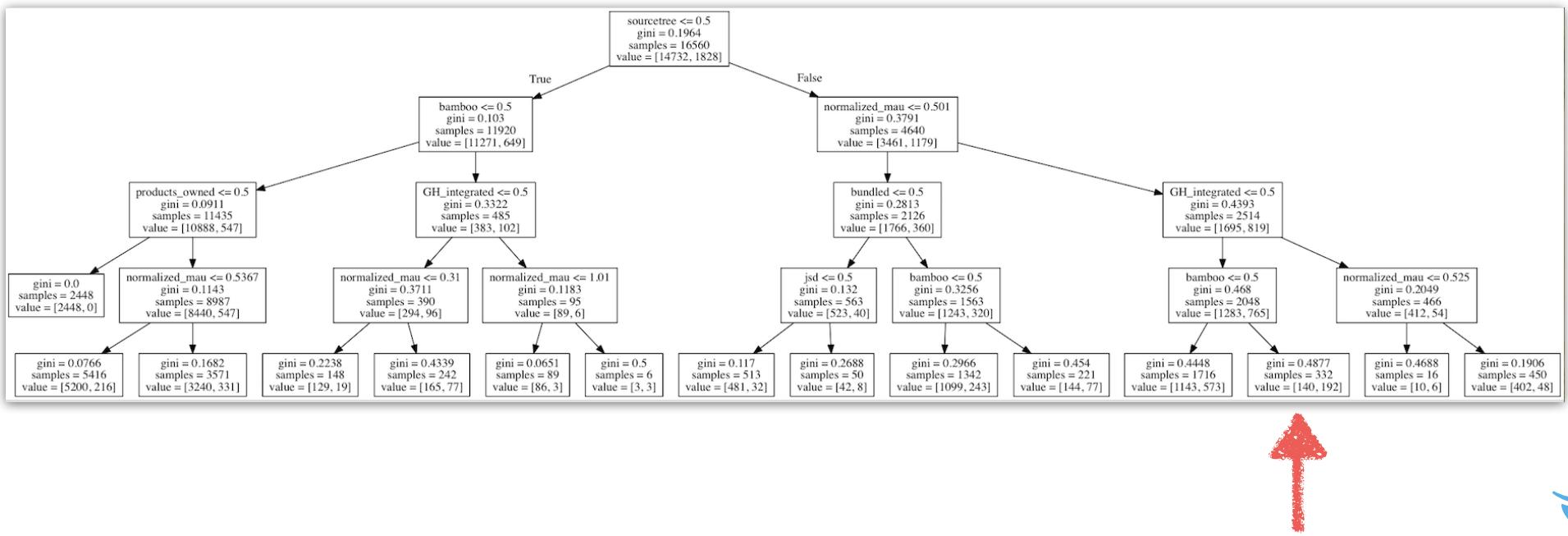


# KNN

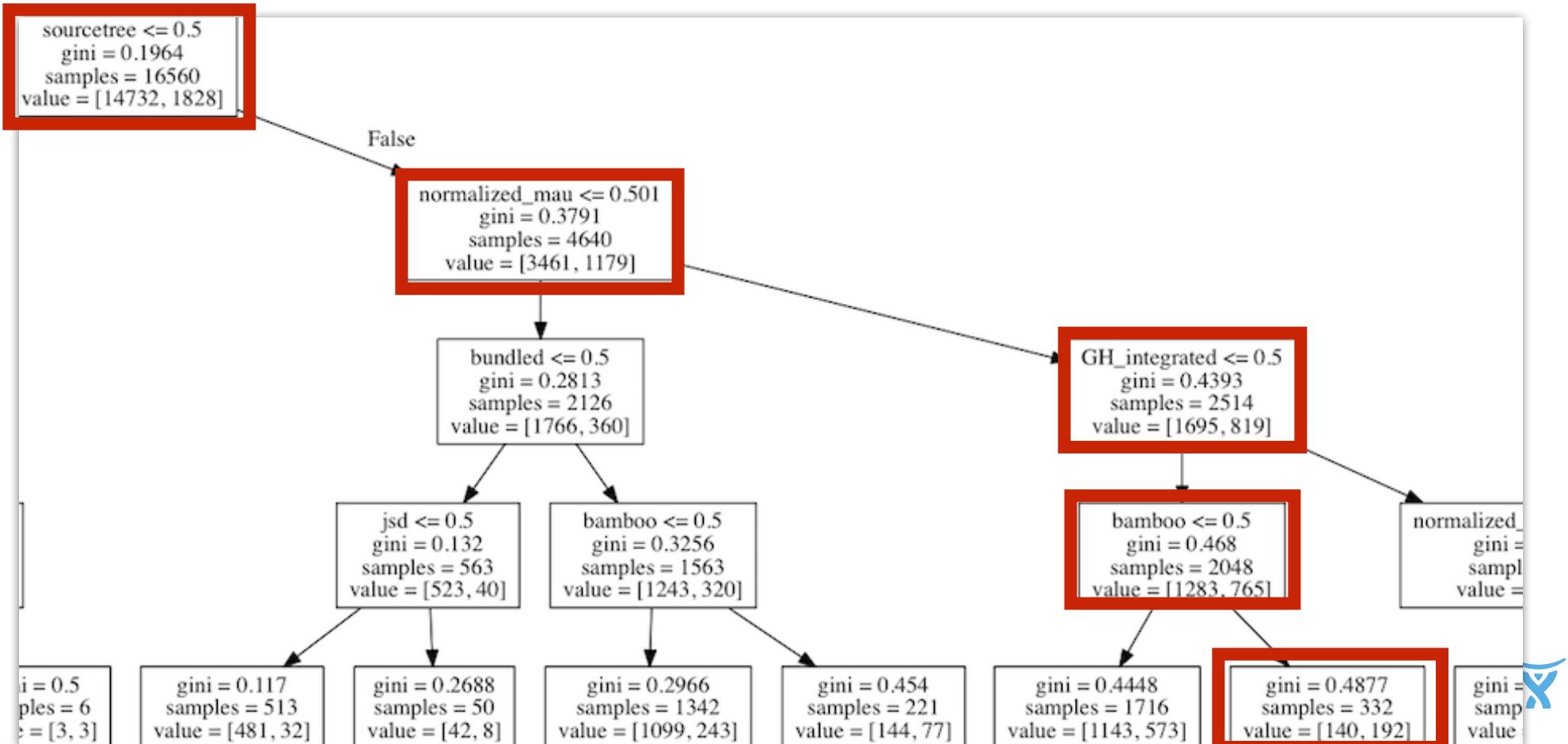
$K = 50$



# Decision Tree



# Decision Tree

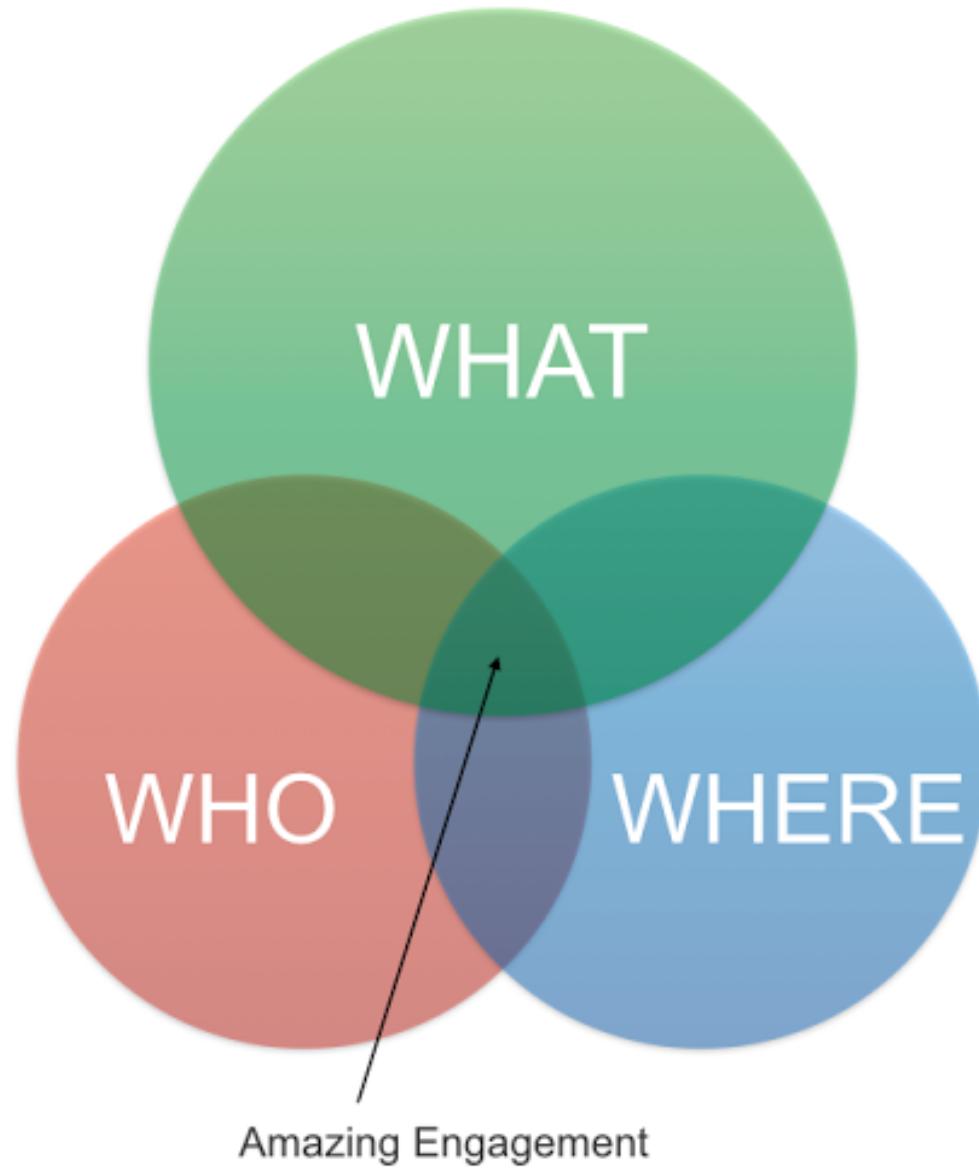


# Decision Tree

```
##Importance
treeclf_features = pd.DataFrame({'feature':feature_cols, 'importance':treeclf.feature_importances_})
treeclf_features.sort_values(['importance'], ascending=[0])
```

	feature	importance
4	sourcetree	0.536929
0	normalized_mau	0.168790
8	bamboo	0.144221
1	GH_integrated	0.121312
2	products_owned	0.028747
16	prod_projectcreated	0.000000
25	prod_kanban	0.000000
24	prod_epics	0.000000





## System Dashboard

Tools

### Introduction



#### Welcome to JIRA

Not sure where to start? Check out the [JIRA 101 guide](#) and [Atlassian training course](#).

You can [customize this text](#) in the Administration section.

### Assigned to Me

You currently have no issues assigned to you. Enjoy your day!

### Activity Stream

#### Your Company JIRA



February 10



**Administrator** started progress on SS-1 - New issue

 10/Feb/16 5:45 PM Comment

February 08



**Administrator** updated the Rank of SPAC-7 - As a user I want to... because it is awesome...

 08/Feb/16 5:20 PM Comment

**Administrator** updated the Sprint of SPAC-7 - As a user I want to... because it is awesome...

 08/Feb/16 5:20 PM Comment

**Administrator** updated the Rank of SPAC-8 - As a user I want to... because it is awesome...

 08/Feb/16 5:15 PM Comment

**Administrator** updated the Rank of SPAC-8 - As a user I want to... because it is awesome...

 08/Feb/16 5:15 PM Comment

### Bitbucket Pipelines NEW



Your team can build, test, and deploy from Bitbucket. With Bitbucket there's no CI server to setup, user management to configure, or repositories to synchronize.

[Sign up for free](#) Close

# Next Steps

- Refine the model
  - Include more in product behavior
- Expand to user level
  - Predict the role of a user based on in product behavior
- JIRA License is based on license today —> would like to change to license size on purchase date





# Thank you!



JESSICA • GROWTH ANALYST • ATLASSIAN