

# Predict Department

-For a given snippet of conversation/summary

-Ramesh Babu DSI-113020

# Problem Statement

In Fortune 500 companies , like paypal having 25,000 employes and having 3000 projects / teams, it is a big challenge for a bug to be navigated to a

1. Right Department ,
2. Right Team ,
3. Right Component,
4. Right developer,
5. With right Severity.

## How serious the impact is

- Lost Revenue
- Will be Out Of SLA (OOSLA)
- Knowledgeable individuals
- Assignee Out of Scope
- Frontline - Standard Operating Procedure (SOP)  
Documents
- Severity Issues

## Who am I,

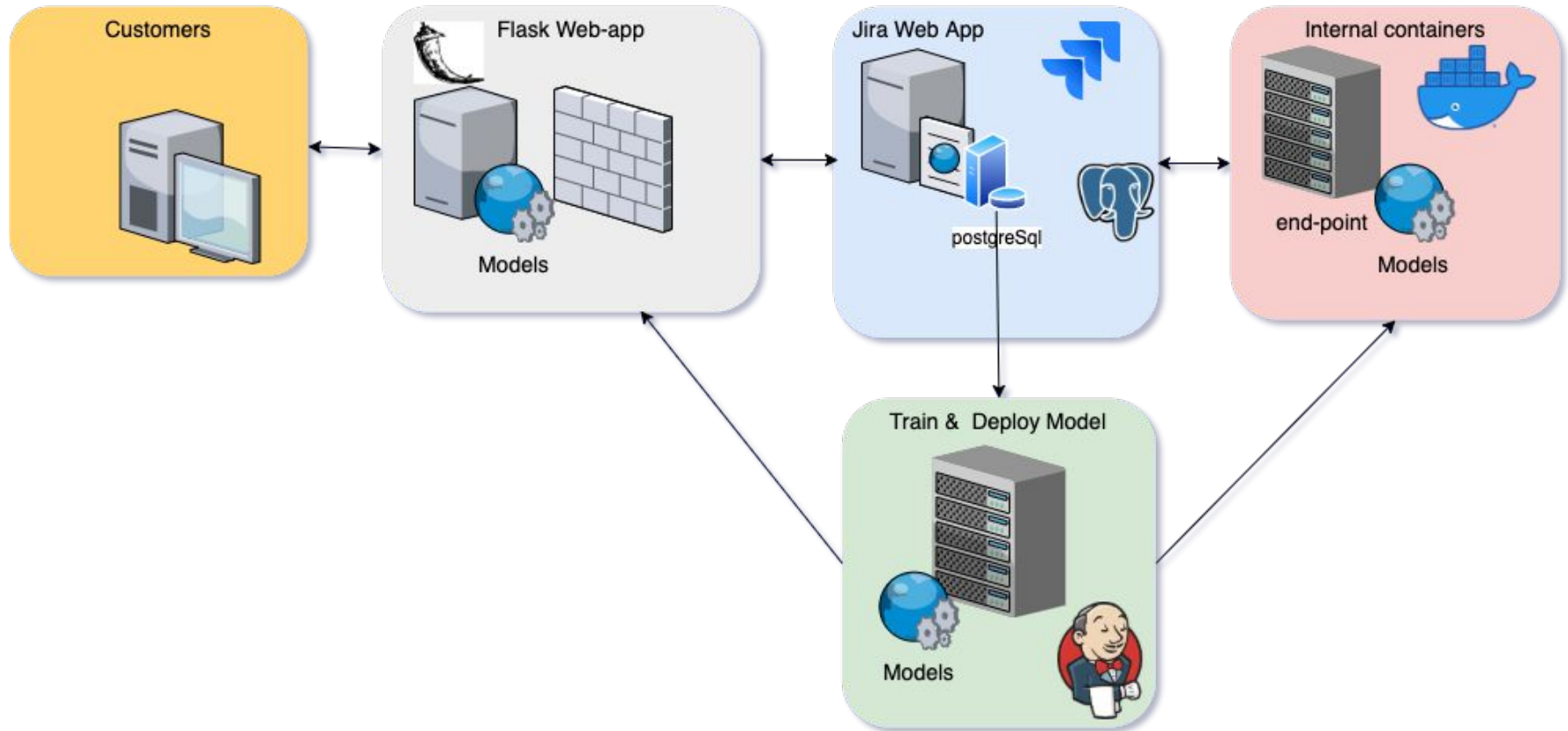
- Handled 3000+ teams across globe
- Team size 10 - 750 members
- Designed and implemented 1000+ workflows, screens and rules
- Trained and Educated 1000s of Trainers ( Train the Trainers)
- I know what is the expectation from top level CTO to the developer working on the code

# What is my solution for this processes

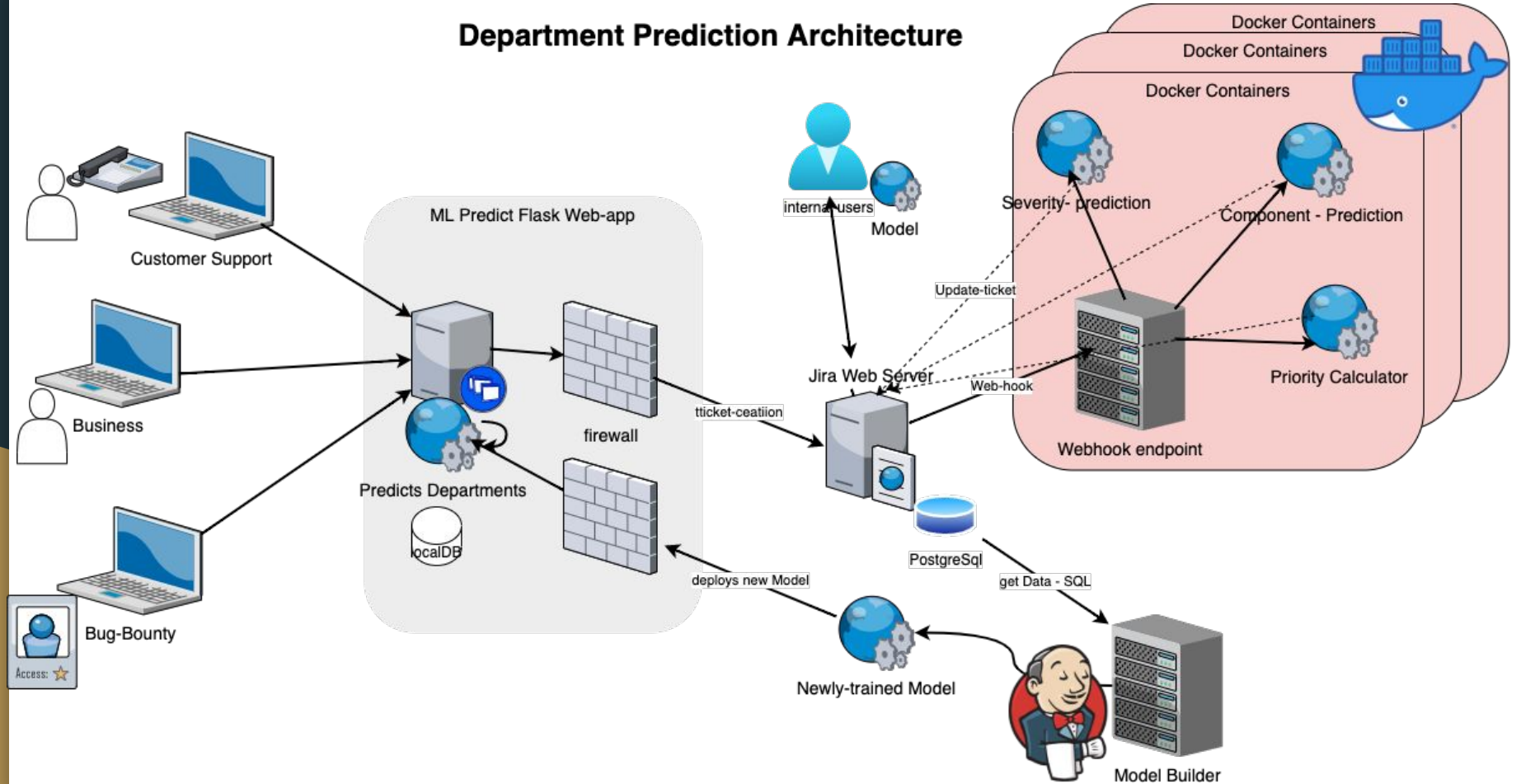
- Drill down multiple ML models ( divide and conquer )
  - Project prediction
  - Within project predict Components
  - With combination of ticket content and Components predict Severity
- Train and improve the efficiency of the model periodically

# Quick Demo

## High Level Architecture Diagram



# Department Prediction Architecture





# Machine Learning Page

This is a Prototype for Machine Learning Algorithm to create ticket in predicted project.

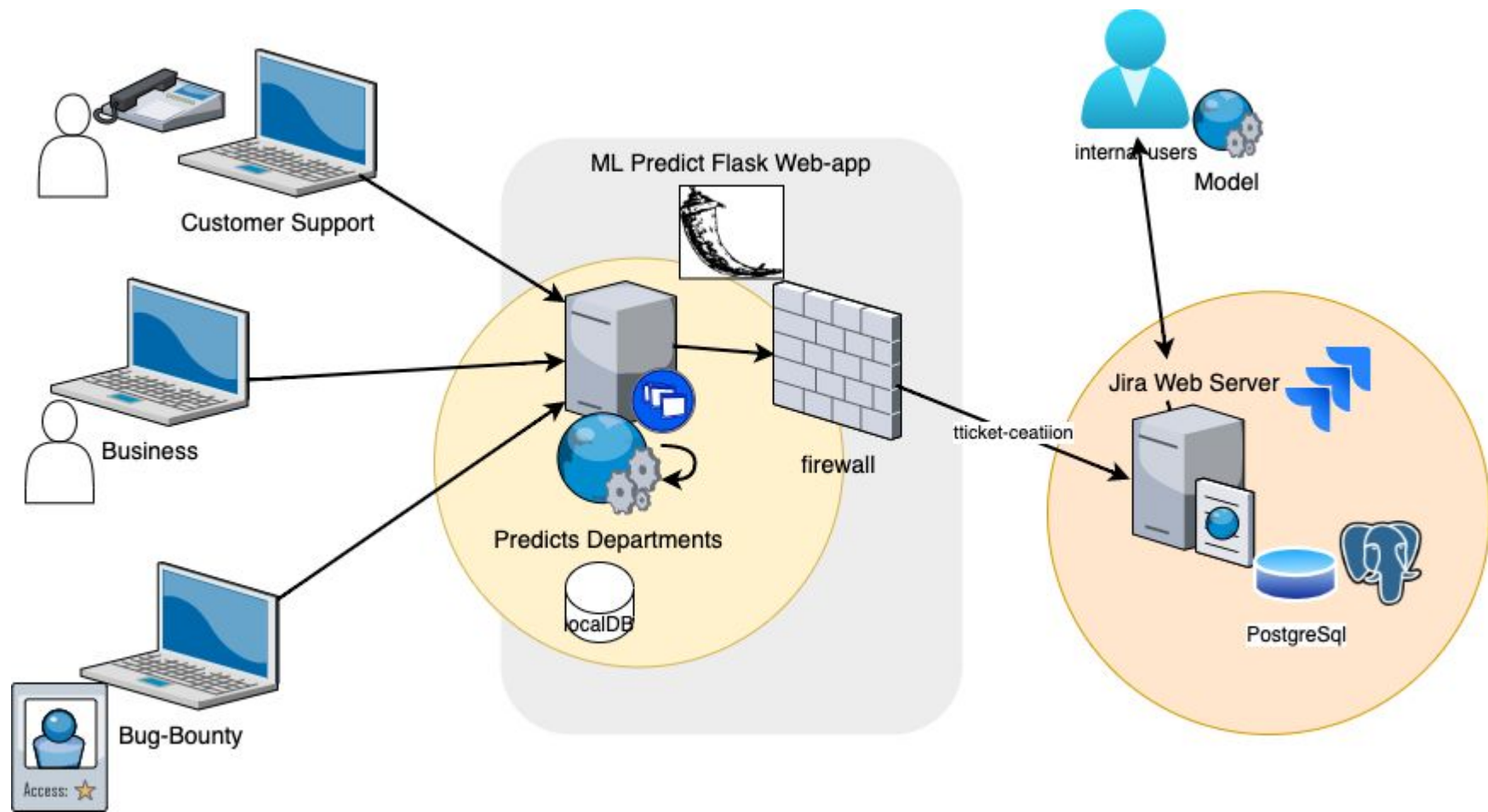
## Create a Ticket in Jira

Please give Summary of the ticket

## Your Previous Tickets

Customer Support ID	Timestamp	Summary	Action
089f4f314246f844ed192f9397f14537b255ce9c	2021-02-27 10:48:11.068524	who is the actor in matrix - * MOV-19587- Clair Hester *	<a href="#">Delete</a>
bbd4e16ec8224cfd5b4688df0b9466fc44c1bdd2	2021-03-01 10:27:53.134074	what went wrong on Jan-26 at capitol building - * PD-10188- None *	<a href="#">Delete</a>
12fa11e838efab891fe3d1e28706bf3203ec3db3	2021-03-01 10:33:04.747579	who is the actor in matrix - * Movies - MOV-19588- Clair Hester *	<a href="#">Delete</a>
4d391dc290cb6bf0ac3a500e99e57359cbe84144	2021-03-01 10:38:49.947862	I died a little After I saw how my mom defrosting meat - * Cooking - CK-19722- Jacob Koehler *	<a href="#">Delete</a>
ceddbb8c34de5891f08fffd0c9cde748b9665801	2021-03-01 11:19:55.650162	I need some time to read - * Books - BK-19607- Ramesh Babu *	<a href="#">Delete</a>
d8a4859f93c40024d03974d8d6de2a19589f827a	2021-03-02 20:27:56.893297	Jane Street Market Prediction Challenge - * Machine Learning - ML-25146- Jeff Hale *	<a href="#">Delete</a>
782ca3b4f046a583494d787c338690f5af6ef4a4	2021-03-02 21:35:01.573056	Is there anyone to help me to make apple pie - * Cooking - CK-19723- Jacob Koehler *	<a href="#">Delete</a>





Political Discussion / PD-10193

## what went wrong on Jan-26 at capitol building



Edit



Comment

Assign

More ▾

To Do

In Progress

Workflow ▾

Admin ▾



Exp

### Details

Type: Bug  
Priority: Medium  
Labels: None  
Reddit ID: lhkv3x  
Reddit url: [https://www.reddit.com/r/PoliticalDiscussion/comments/flss8m/why\\_are\\_illegal\\_immigrants\\_counted\\_for\\_district/](https://www.reddit.com/r/PoliticalDiscussion/comments/flss8m/why_are_illegal_immigrants_counted_for_district/)  
created\_utc: 1613647784  
Severity: S3

Status: **TO DO** (View Workflow)  
Resolution: Unresolved

### Description

How to Write Effective Tickets  
Whenever our developers review their tasks in Jira they look for details; the key is to include the right amount of details without overloading the ticket with excess information. The main purpose of tickets in Jira is to communicate with developers, and good communication minimizes project time, prevents confusion, and are a great reference source during the review phase.

### Attachments

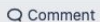


Drop files to attach, or [browse](#).

### Activity

All Comments Work Log History Activity

There are no comments yet on this issue.



Comment

### People

Assignee: Chuck D  
[Assign to me](#)  
Reporter: admin  
Votes: 0  
Watchers: 1 [Stop watching this issue](#)

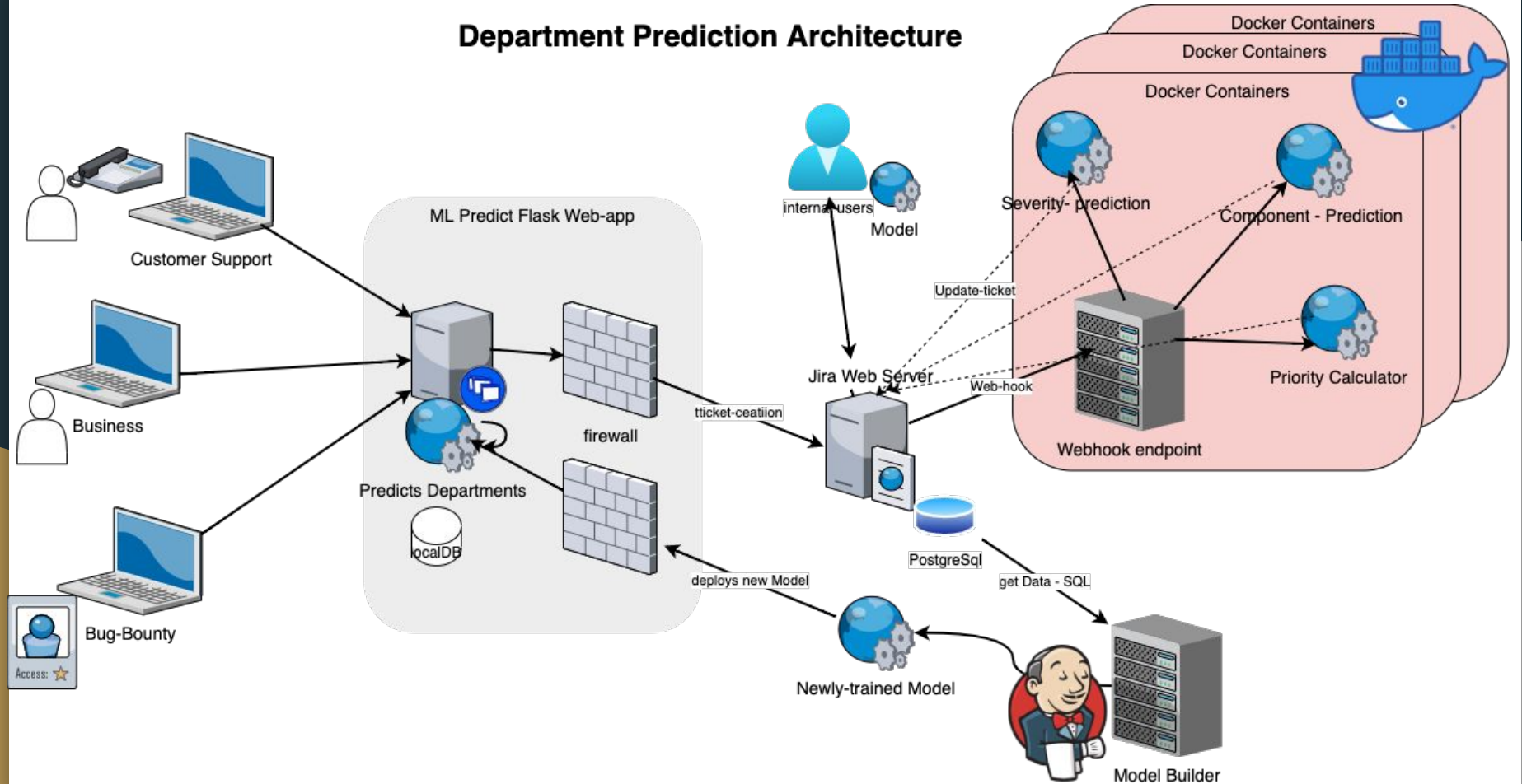
### Dates

Created: 5 minutes ago  
Updated: 3 minutes ago  
created\_gmt: 21/Feb/21 6:02 PM

### Agile

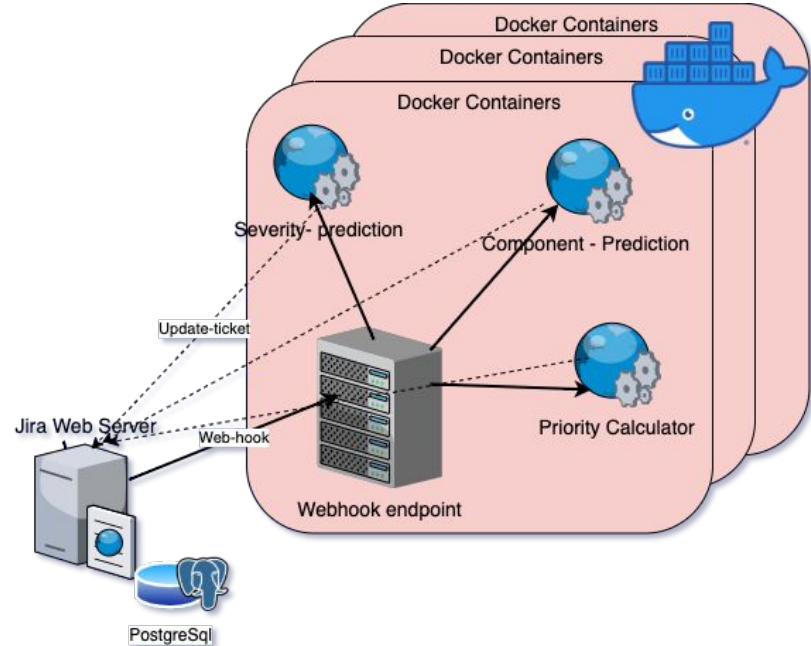
[View on Board](#)

# Department Prediction Architecture

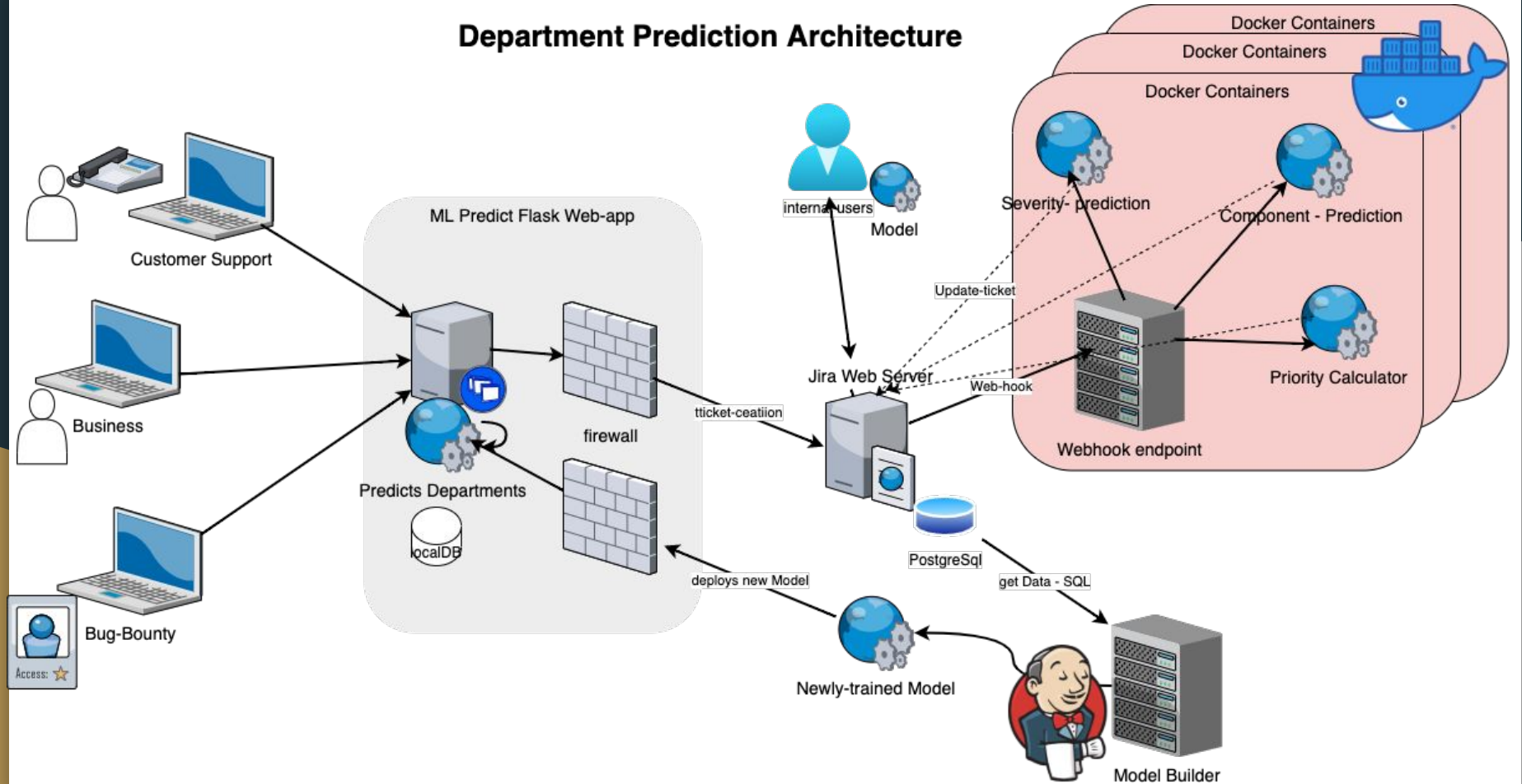


# Background processes

1. Web-hook will be fired through Jira to wake up our Project level model
2. Say Project ML model does NLP and will assign
  - a. Project Component ( Database, Network, Frontend-GUI, Backend-Dev, Architect)
  - b. Component Lead ( will be provided by Jira )
  - c. Severity ( S1..S5)
  - d. Priority (P0..P4)
  - e. Due Date ( calculated)



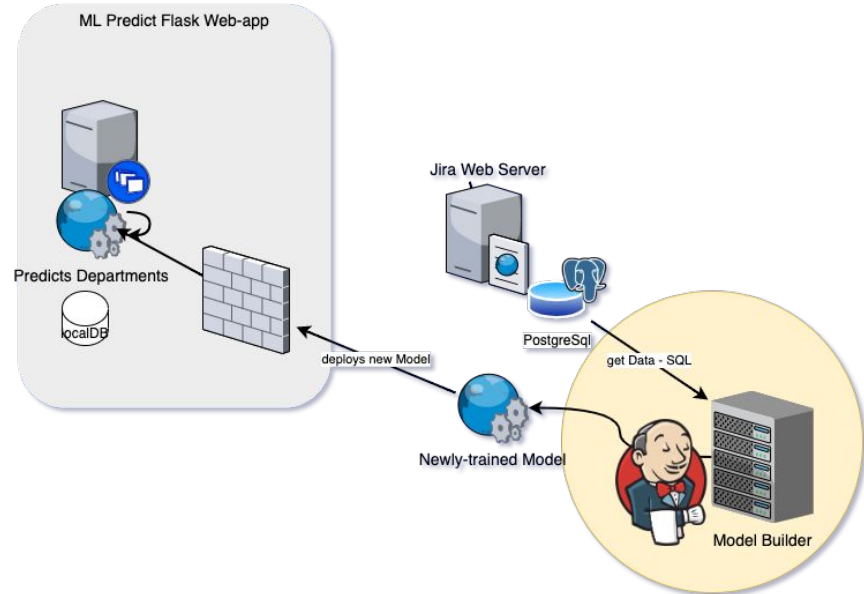
# Department Prediction Architecture



# Train Model with Delta

Every evening , our model has to go through

1. Re-train the delta data from the last model
2. Re-train with wrong Predictions
3. Send out statistics about the model to Data-scientist
  - a. Confusion Matrix
  - b. Train and Test Scores
  - c. ROC curve





## Projects

Cre

Q Search...

Project ↑	Key	Project type	Project lead	Project category	Last issue update	Issues	Actions
 Books	BK		Ramesh Babu	Machine Learning ...	22 Feb 2021	19,605	...
 Cooking	CK		Jacob Koehler	Machine Learning ...	24 Feb 2021	19,718	...
 Machine Learning	ML		Jeff Hale	Machine Learning ...	25 Feb 2021	25,145	...
 Movies	MOV		Clair Hester	Machine Learning ...	22 Feb 2021	19,585	...
 Political Discussion	PD		Chuck D	Machine Learning ...	21 Feb 2021	10,185	...
 Testing Project	TP		Ramesh Babu	Machine Learning ...	21 Feb 2021	16	...



## Components



QUICK FILTERS:

Active

Archived

Add

Component	Status	Issues	Lead	Description	Default assignee	Actions
Backend Development	ACTIVE	0 Issues	Jeff Hale		Component lead	...
Database	ACTIVE	1 Issue	Jacob Koehler		Component lead	...
Frontend GUI	ACTIVE	0 Issues	Clair Hester		Component lead	...
Network	ACTIVE	0 Issues	Chuck D		Component lead	...



Machine Learning / ML-25143

## [N] Kaggle/Jane Street Market Prediction Challenge



Comment

Assign

More

To Do

In Progress

Workflow

Admin



Export

### Details

Type: **Bug** Status: **TO DO** (View Workflow)  
Priority: **Highest** Resolution: Unresolved  
Component/s: **Database**  
Labels: None  
Reddit ID: jzndmq  
Reddit url: [https://www.reddit.com/r/MachineLearning/comments/jzndmq/n\\_kagglejane\\_street\\_market\\_prediction\\_challenge/](https://www.reddit.com/r/MachineLearning/comments/jzndmq/n_kagglejane_street_market_prediction_challenge/)  
created\_utc: 1606156330  
Severity: **S3**

### Description

<https://www.kaggle.com/c/jane-street-market-prediction/overview> (<https://www.kaggle.com/c/jane-street-market-prediction/overview>)

&#x200B;

>\*\*\*"Buy low, sell high." It sounds so easy....\*\*

>

>In reality, trading for profit has always been a difficult problem to solve, even more so in today's fast-moving and complex financial markets. Electronic trading allows for thousands of transactions to occur within a fraction of a second, resulting in nearly unlimited opportunities to potentially find and take advantage of price differences in real time.

>

>In a perfectly efficient market, buyers and sellers would have all the agency and information needed to make rational trading decisions. As a result, products would always remain at their "fair values" and never be undervalued or overpriced. However, financial markets are not perfectly efficient in the real world for a number of reasons.

>

### People

Assignee: Jeff Hale  
[Assign to me](#)  
Reporter: Ramesh Babu  
Votes: **0** [Vote for this issue](#)  
Watchers: **1** [Stop watching this issue](#)

### Dates

Due: **27/Feb/21**  
Created: 2 days ago  
Updated: Just now  
created\_gmt: 20/Nov/20 1:11 PM

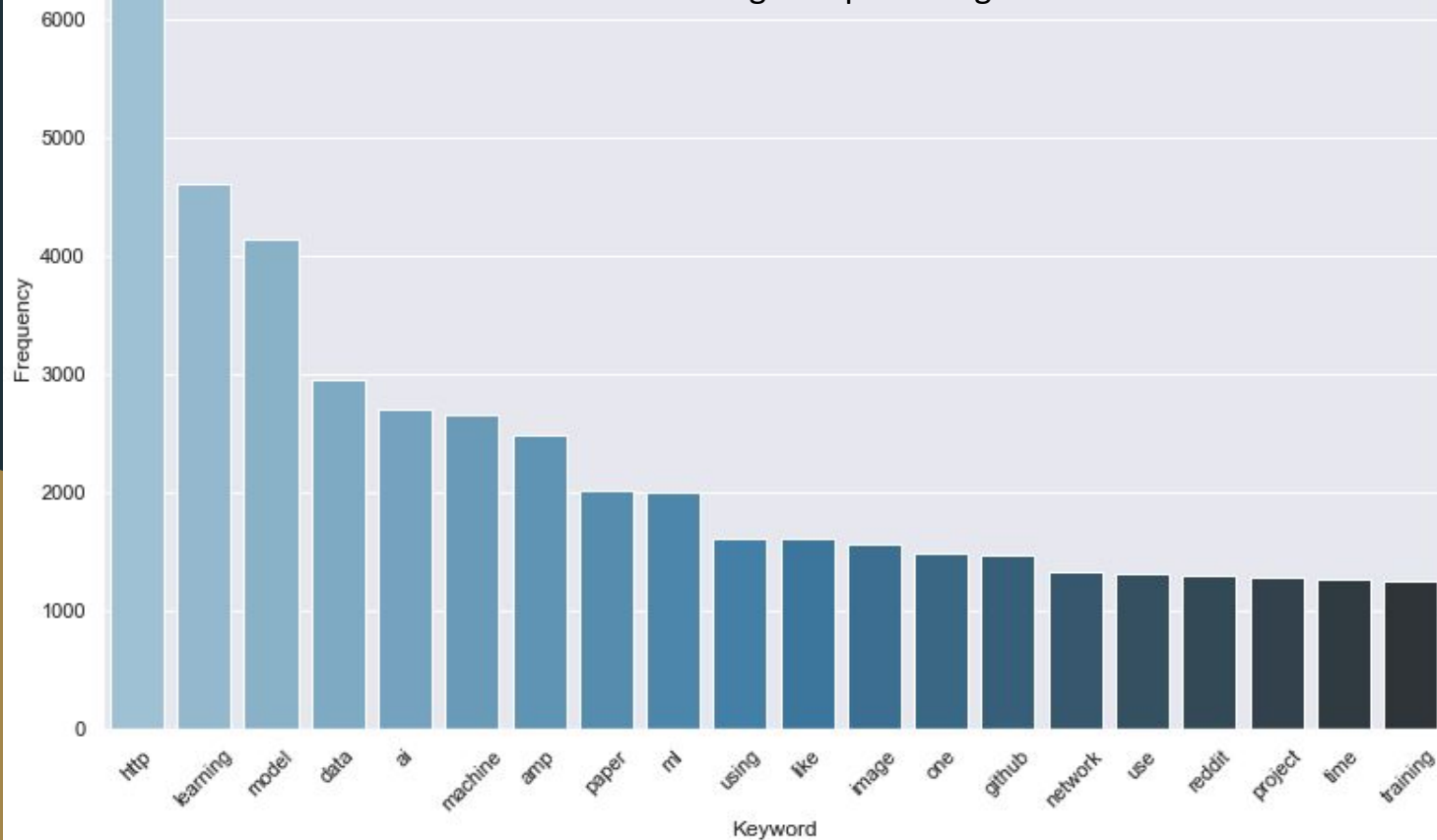
### Time Tracking

Estimated: 2d  
Remaining: 2d  
Logged: Not Specified

### Agile

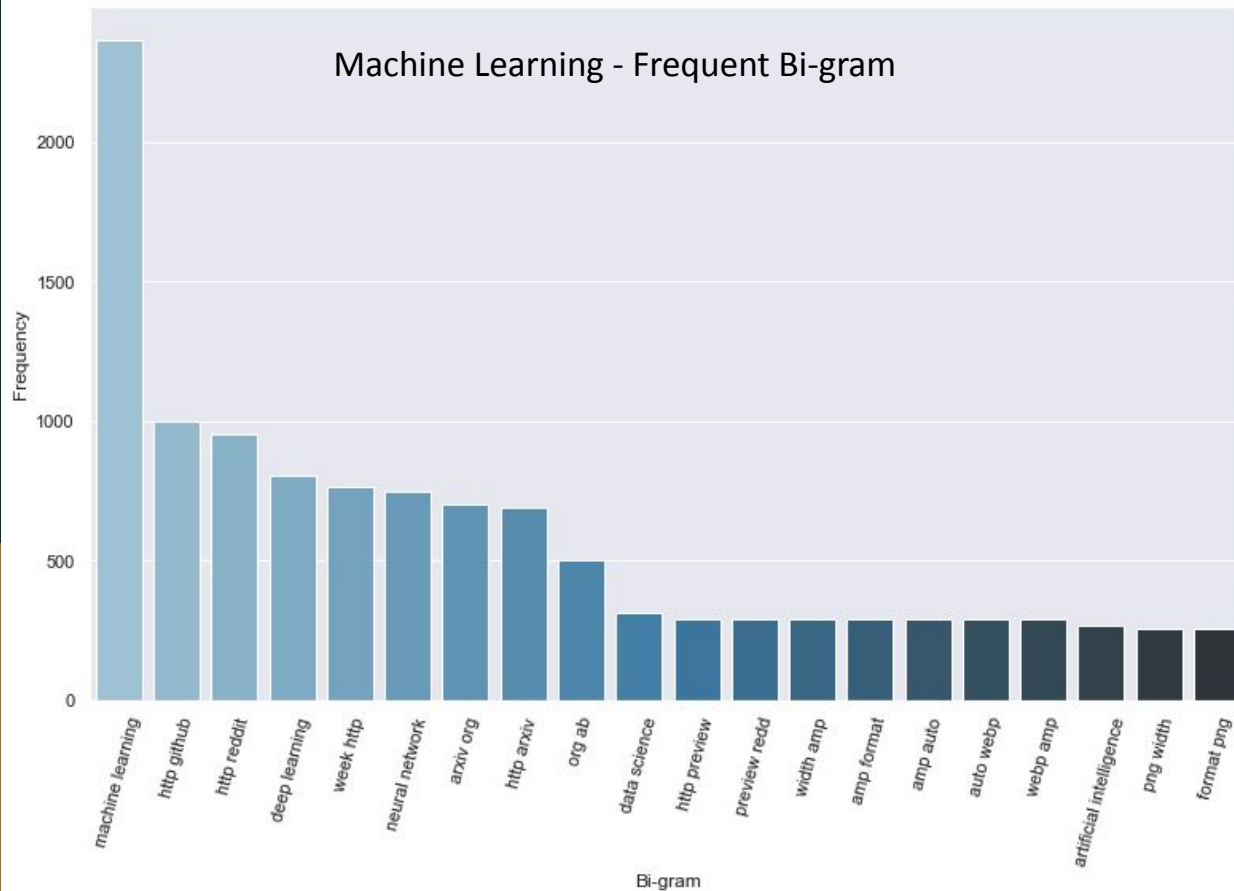
[View on Board](#)

# Machine Learning- Frequent Single Word



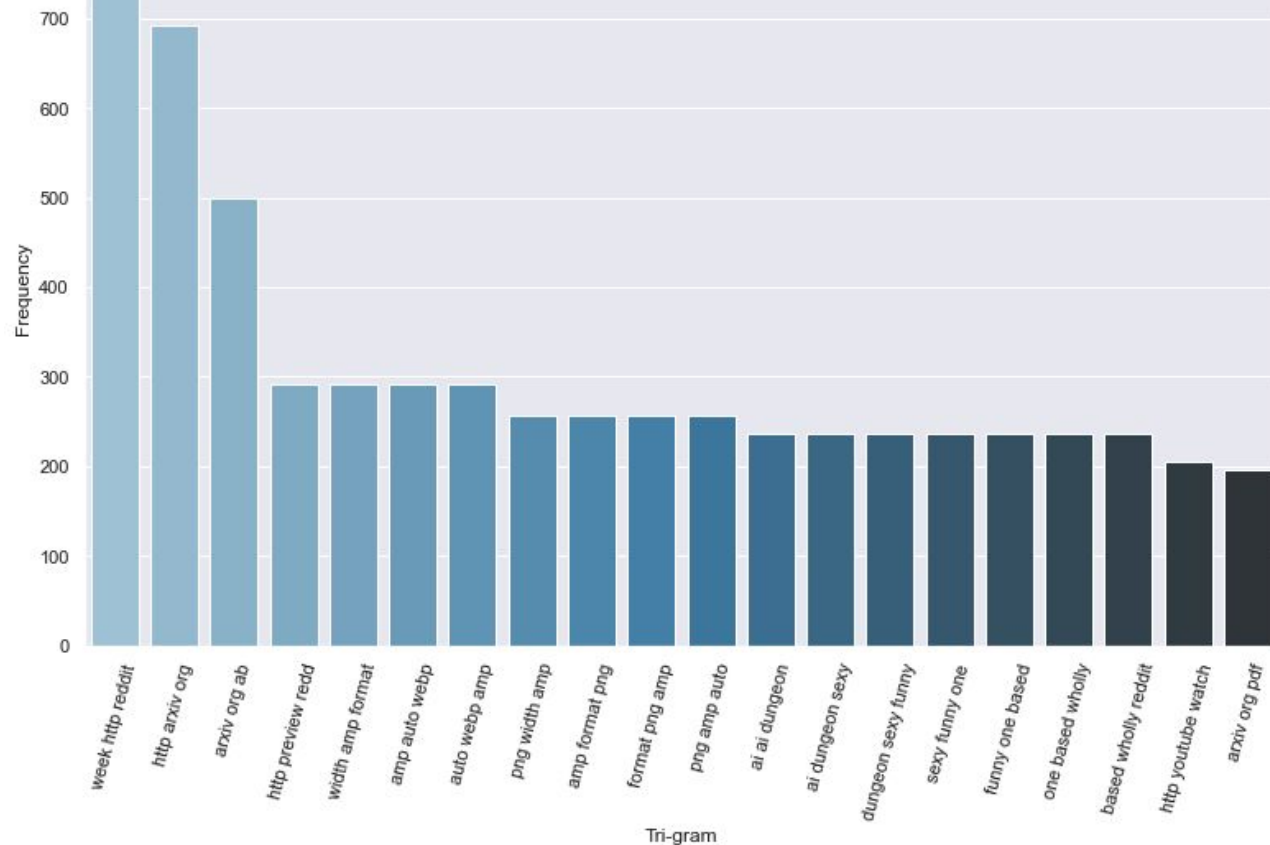
Keyword	Frequency
http	6474
learning	4611
model	4139
data	2944
ai	2696
machine	2648
amp	2477
paper	2011
ml	2003
using	1613
like	1601
image	1565
one	1475
github	1470
network	1330
use	1303
reddit	1293
project	1271
time	1266
training	1249

## Machine Learning - Frequent Bi-gram



Bi-gram	Frequency
machine learning	2366
http github	997
http reddit	951
deep learning	807
week http	762
neural network	749
arxiv org	701
http arxiv	693
org ab	500
data science	311
http preview	292
preview redd	292
width amp	292
amp format	292
amp auto	292
auto webp	292
webp amp	292
artificial intelligence	266
png width	257
format png	257

## Machine Learning - Frequent Tri-gram



Tri-gram	Frequency
week http reddit	761
http arxiv org	693
arxiv org ab	500
http preview redd	292
width amp format	292
amp auto webp	292
auto webp amp	292
png width amp	257
amp format png	257
format png amp	257
png amp auto	257
ai ai dungeon	236
ai dungeon sexy	236
dungeon sexy funny	236
sexy funny one	236
funny one based	236
one based wholly	236
based wholly reddit	236
http youtube watch	205
arxiv org pdf	196

# What did we gain so far..

- Eliminated Triage Team completely
- Saved some time for Project Leads
- Frictionless backlog grooming
- Severity S1-S2 injected into Agile Board
- Bottle-neck removed & Bugs life has been shortened
- Standard Operating Procedure ( SOP) document minimised



How did I achieve this using ML

# Regular Machine Learning Models

- Logistic Regression
- MultinomialNB
- RandomForest
- ExtraTrees
- K-NearestNeighbors
- SVC
- AdaBoostClassifier
- GradientBoostingClassifier



## Neural Network

- KerasClassifier (dropout:0.5, reg:l2, units:64)
- CNN
- RNN
  - SimpleRNN
  - GRU
  - LSTM

And the best model is.....

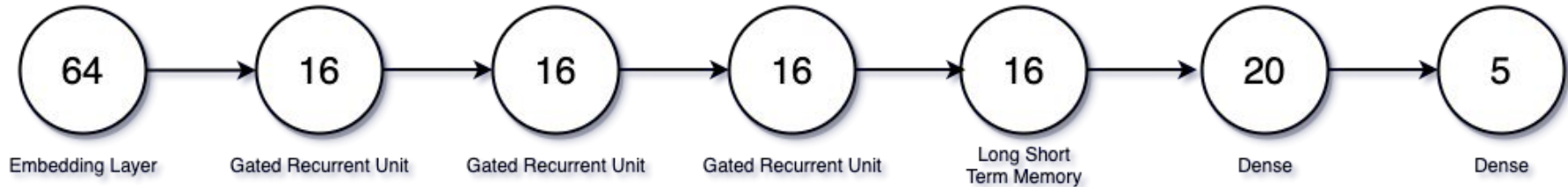
Recurrent Neural Network (RNN)

Gated Recurrent Unit (GRU)

Long Short Term Memory (LSTM)

# Neural Network Diagram

**Under the Hood**  
A visual representation of Neural Network



Model: "sequential\_2"

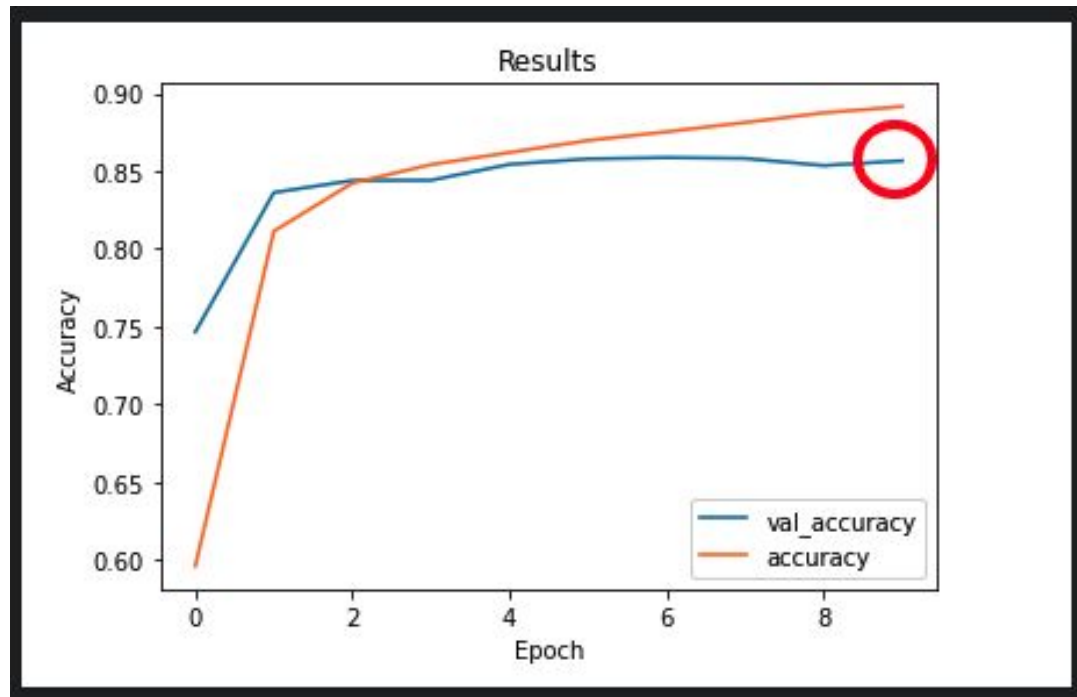
Layer (type)	Output Shape	Param #
embedding_2 (Embedding)	(None, None, 64)	64000
gru_6 (GRU)	(None, None, 16)	3936
gru_7 (GRU)	(None, None, 16)	1632
gru_8 (GRU)	(None, None, 16)	1632
lstm_2 (LSTM)	(None, 16)	2112
dense_2 (Dense)	(None, 20)	340
dense_3 (Dense)	(None, 5)	105

Total params: 73,757

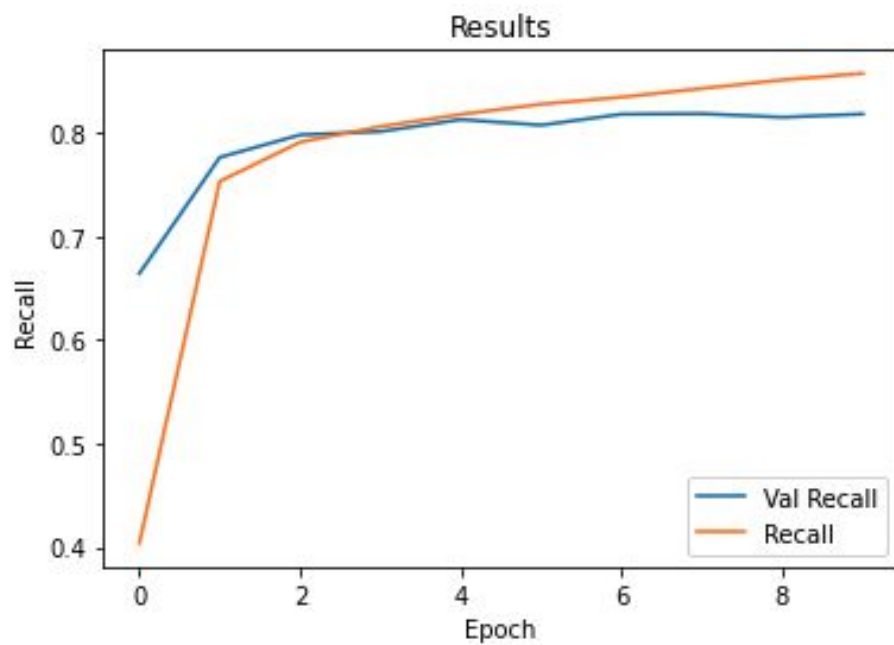
Trainable params: 73,757

Non-trainable params: 0

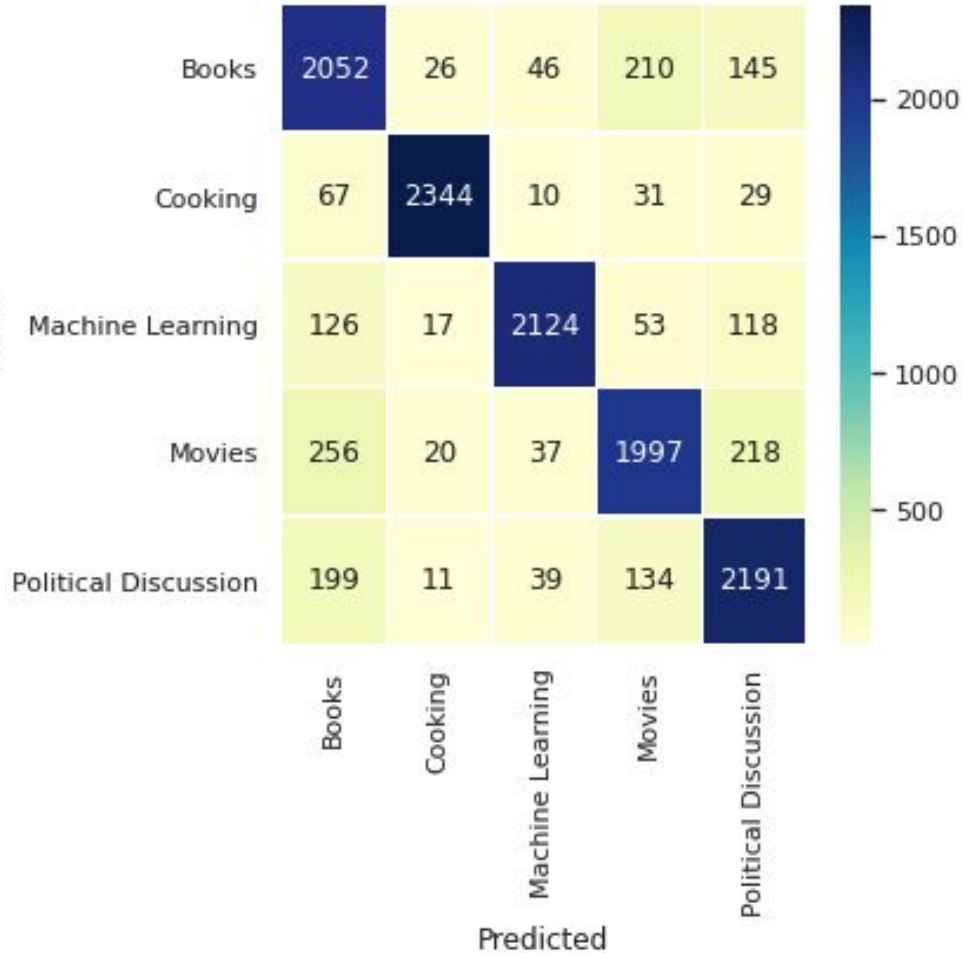
Train accuracy: 0.8566



Train Recall :0.8188



Actual



Confusion Matrix

	precision	recall	f1-score	support
Books	0.76	0.83	0.79	2479
Cooking	0.97	0.94	0.96	2481
Machine Learning	0.94	0.87	0.90	2438
Movies	0.82	0.79	0.81	2528
Political Discussion	0.81	0.85	0.83	2574
accuracy			0.86	12500
macro avg	0.86	0.86	0.86	12500
weighted avg	0.86	0.86	0.86	12500



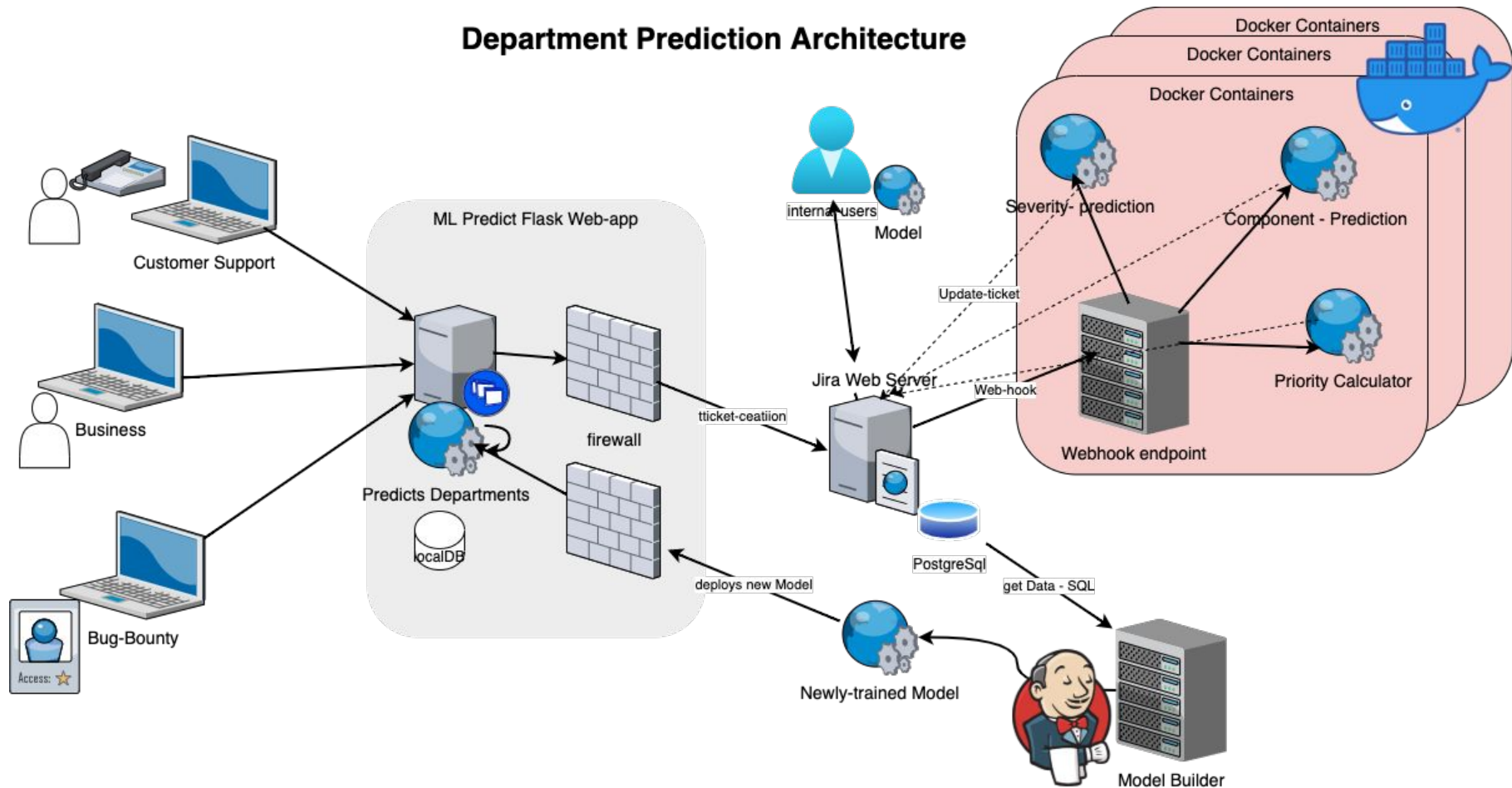
# Future Implementation - Replace Frontline with AI

**Use-Case :** There is company wide policy to reset the password every 90 days. Most of the users forget on 91st day they has changed the password and usually try with old password and they system get locked

**Solution:** we can start creating `Question - Answer` machine learning algorithm to ask questions in Slack channel installed in user phone .

Questions ...?

## Department Prediction Architecture



### **Demo Conversation:**

User: Hey @FL-bot , I got locked out, please unlock me

FL-bot: Yes, I see you have been locked out, by trying to login with wrong password from this IP from San Jose

FL-bot: I have texted you a code to your registered phone number under Paypal

User: 3452-7827-9844

FL-bot: Thanks, .. Let me unlock your account

FL-bot: I see you have been successful logged in to paypal system from this IP from San Jose.

FL-bot: Please rate my service, so that I can train myself for future requests

User: Satisfied, five star !!

User: @jira-bot, I am not able to reach jira from `ete01-server.paypal.com` through restapi

Jira-bot: Can you please try to do `telnet jira-server 8080`

User: here are the output

Trying 192.168.2.290...

Connected to jira-server.

Escape character is '^['

Jira-bot: Looks like there are no firewall issues from jira server to your server

jira-bot : please try..curl -u userid@paypal.com:password -X GET -H "Content-Type: application/json"  
<https://jira.paypal.com/rest/api/2/issue/createmeta>

User: here is the output `Please check your credentials`

Jira-bot: Looks like your password has been expired, please reset your password