

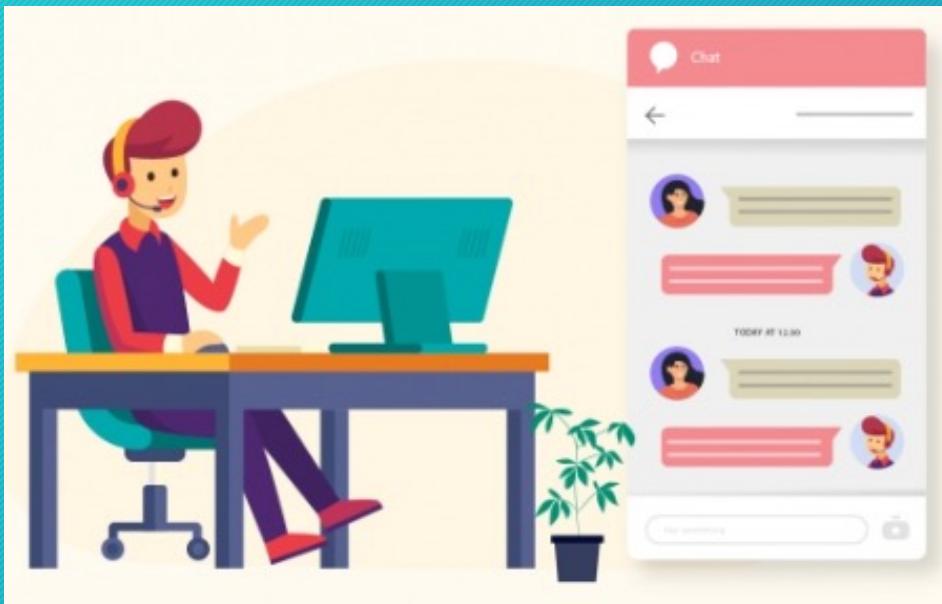
City of El Paso's Ticketing System Classification using NLP methods

Outline

- Background
- Problem
- Exploratory data analysis
- Neural network model
- Results & conclusion
- Future work

Background

- City Departments such as Fire Department or Libraries can use ticketing system to request help to IT department. Examples include trouble shooting or equipment replacement
- IT expert reads request and manually assigns to correct IT team
- Want to create a model that classifies tickets to save IT expert some time



A screenshot of a ticketing system interface. The top navigation bar includes 'Contacts', 'Conversations', 'Marketing', 'Sales', 'Service', 'Automation', and 'Dashboards'. The main area is titled 'Tickets' and shows a list of 22 tickets. The columns are: TICKET NAME, TICKET STATUS, CREATE DATE, TICKET PRIORITY, HUBSPOT OWNER, and DATE OF LAST ENGAGEMENT. The table lists various tickets with details such as 'Error Message' (New, Apr 9, 2018, High, Pat Participant (cloudco...), Apr 4, 2018), 'Help adding a user' (New, Apr 9, 2018, High, Pat Participant (cloudco...)), and 'I just deleted all my contacts. HELP!!' (Waiting on contact, Apr 5, 2018, High, Sophie Higgs (shiggs@...)).

TICKET NAME	TICKET STATUS	CREATE DATE	TICKET PRIORITY	HUBSPOT OWNER	DATE OF LAST ENGAGEMENT
Error Message	New	Apr 9, 2018	High	Pat Participant (cloudco...)	
Help adding a user	New	Apr 9, 2018	High	Pat Participant (cloudco...)	
Issue updating attendee	New	Apr 4, 2018	Low	Pat Participant (cloudco...)	Apr 4, 2018
I just deleted all my contacts. HELP!!	Waiting on contact	Apr 5, 2018	High	Sophie Higgs (shiggs@...)	
Invoice issue	Waiting on us	Apr 5, 2018	High	Sophie Higgs (shiggs@...)	
Export doesn't contain the right fie...	Waiting on contact	Apr 5, 2018	Low	Sophie Higgs (shiggs@...)	
Billing page bug	New	Apr 2, 2018	High	Sophie Higgs (shiggs@...)	
Free trial?	New	Apr 6, 2018	High	Jenna Glat (jenna.glat@...)	
Beta program	New	Mar 30, 2018	High	Jenna Glat (jenna.glat@...)	
Help finding Settings	New	Mar 28, 2018	High	Jenna Glat (jenna.glat@...)	
Help using new feature	New	Mar 27, 2018	High	Jenna Glat (jenna.glat@...)	
Lost item	New	Apr 5, 2018	High	Laura Mikulay (lauratest...)	
Credit card issue	Waiting on us	Apr 4, 2018	High	Laura Mikulay (lauratest...)	
Storage question	Waiting on us	Apr 2, 2018	Low	Laura Mikulay (lauratest...)	
test	New	Apr 25, 2018	Low	Unassigned	
new test ticket	New	Apr 18, 2018	Low	Unassigned	
Extension of free trial?	New	Apr 9, 2018	Low	Unassigned	
My account breaks when I try to up...	New	Apr 5, 2018	Low	Unassigned	

- Can we classify incoming tickets and assign them to the correct support IT team using Natural Language Processing?

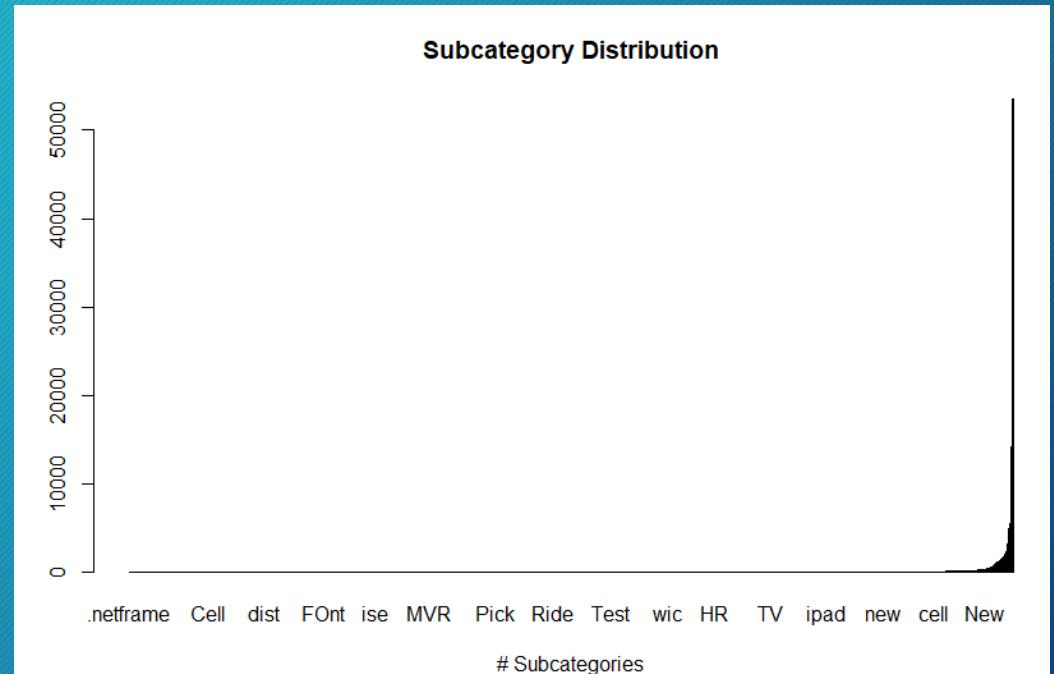
Datasets origin

- Cherwell ticketing system dataset:
Provided by City IT, no private fields to preserve Privacy.

Incident ID	Priority	Description	Customer Department	Owned By Team	Service	Category	Subcategory
472861	3	PLEASE NOTE THAT WE'RE ADDING NUMBERS ON RECOVERIES. CAN WE ADD A COLUMN City Manager	Information Technology	Web Administration	Web Services	Information Update	Content
472862	3	Work on Teams virtual backgrounds. Finish out Phase 1 timeline for SilverStripe city \r\n\r\nInformation Technology	Information Technology	Web Administration	Web Services	Information Update	Content
472863	3	Assistance with reinstalling e-mail archives.\r\n\r\nMiranda, Luis	Public Health	Help Desk	Remote Support Services	Software Support	Troubleshooting
472864	3	OMB review meeting of 3 pending encumbrance issues.\r\n\r\nTaking over the last 4	Information Technology	Applications	Enterprise Apps	PeopleSoft Financials	Technical Support
472865	5	Please assist with configuring the new computers at George Perry testing site for the	Information Technology	Network Infrastructure	Network Services	Network Connectivity	Modify Service
472866	2	Teams call: plan for new featureclasses from testing database, new feature services, Information Technology	Information Technology	GIS	GIS Services	Web Viewer	Update
472867	3	Microsoft SQL 2008 server migration to Microsoft SQL 2012.\r\n\r\n\r\nSSRS ACCELA cor	Information Technology	Applications	Enterprise Apps	Accela	Accela Technical Support
472868	2	Teams call: new web map for covid dashboard	Information Technology	GIS	GIS Services	Web Viewer	Update
472869	5	Meeting with the Zoo to assist with failed survey\r\n\r\nAssiting Angel with research	Information Technology	Applications	IT General	General	General
472870	3	Current Laptop: Dell XPS 9350\r\n\r\nService Tag:F7KBPFF2\r\n\r\nCity Tag: 2395550\r\n\r\nDepart	Information Technology	Support Services	Onsite Support Services	Computer/Tablets	Setup/Relocation
472871	3	Issue: Phoenix Expungement - receives unhandled error. See attachment.\r\n\r\nCaller: Police	Information Technology	Applications	Non-Enterprise Apps	Application Support	Technical Support
472873	4	User reporting not able to print from FS19-02	Fire	Support Services	Remote Support Services	Software Support	Papercut
472874	5	Please assist with printer at George perry testing site.	Information Technology	Network Infrastructure	Network Services	Network Connectivity	Modify Service
472875	2	[Performance Office - The 500]\r\n\r\nGood afternoon team, Could you please include th	Information Technology	Web Administration	Web Services	Electronic Form	Form Modification
472876	3	Research PS Script to find H drive folder size	Information Technology	Support Services	Remote Support Services	Software Support	Troubleshooting
472878	3	post meeting notices	Municipal Court	Web Administration	Web Services	Information Update	Content

Exploratory data analysis (EDA)

- Original dataset:
- > glimpse(df)
Rows: 230,805
columns: 23
- Predictor variable is Description and Target variable is Subcategory.
- Unique values for Subcategory: 2146.



Exploratory data analysis (EDA) - Cont.

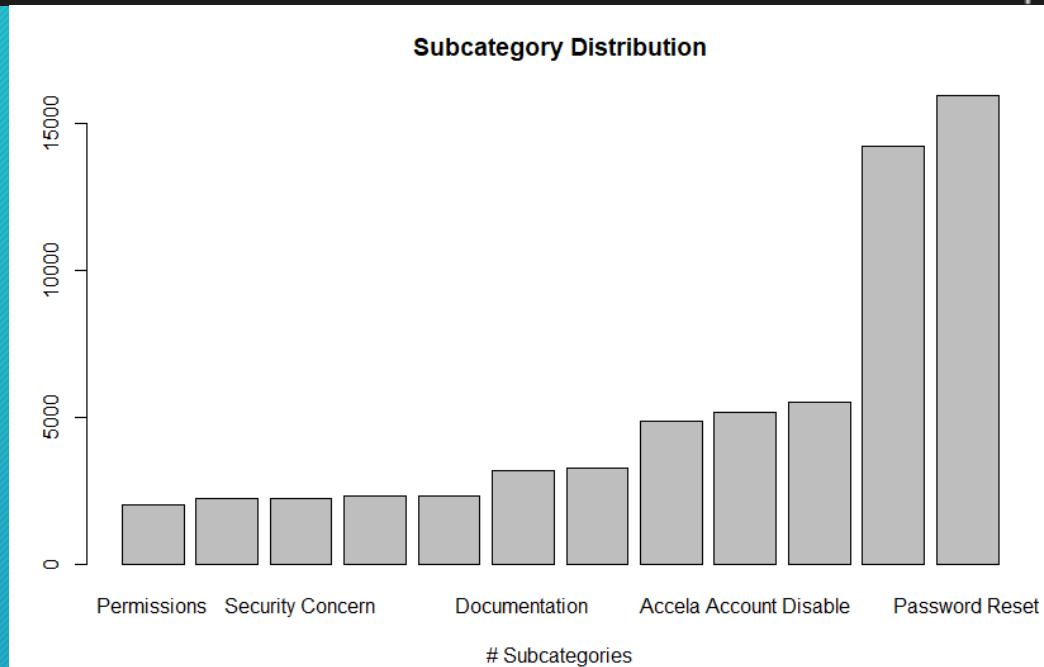
```

Accela Account Disable      5182
Documentation               3204
Permissions                 2021
> unique(df2$subcategory)
[1] "Password Reset"
[5] "Content"
[9] "Creation/Modification"

```

Subcategory	Count	Subcategory	Count	Subcategory	Count	Subcategory	Count
Badge Access	4859	Content	5497	Creation/Modification	2343	Employee Separation	14223
Documentation		Installation/Modification	3271	Password Reset	15952	Permissions	2241
Permissions		Service Down	2340	Unlock Account	2223	Accela Account Disable	
Accela Account Disable						Employee Separation	
						Badge Access	
						Documentation	
						Permissions	
						Security Concern	
						Accela Account Disable	
						Employee Separation	
						Badge Access	
						Documentation	
						Permissions	
						Security Concern	
						Accela Account Disable	
						Employee Separation	

12 most common Subcategory Values:



Neural network model

Target variable, Subcategory, needed to be converted to numeric value and represented as a matrix-like data structure to use it as output nodes in the NN model.

> y-test	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]
[1,]	0	0	0	0	0	0	0	1	0	0	0	0
[2,]	0	0	0	0	0	0	0	1	0	0	0	0
[3,]	0	0	0	0	0	0	1	0	0	0	0	0
[4,]	0	0	0	0	0	0	0	0	0	0	1	0
[5,]	0	0	0	0	0	0	0	0	0	1	0	0
[6,]	0	0	0	0	0	0	0	1	0	0	0	0
[7,]	0	0	0	0	0	0	1	0	0	0	0	0
[8,]	0	0	1	0	0	0	0	0	0	0	0	0
[9,]	0	0	0	0	0	0	0	1	0	0	0	0
[10,]	0	0	0	0	0	0	1	0	0	0	0	0
[11,]	0	0	1	0	0	0	0	0	0	0	0	0
[12,]	0	0	0	0	0	0	0	1	0	0	0	0
[13,]	0	0	0	0	0	0	0	1	0	0	0	0
[14,]	0	0	1	0	0	0	0	0	0	0	0	0
[15,]	0	0	0	0	0	0	0	0	0	0	0	1
[16,]	0	0	0	0	0	0	0	1	0	0	0	0

Neural network model - Cont.

9

Predictor variable Description Remarks:

- Free-text value input by user.
- Removed stop words, e.g., “the”. “a”, “is”, “are”.
- Identified top 1000 most used words
 - “Please **reset** my **password**, I don’t remember it” and “My **password** has expired, I need to **reset** it”
 - “My computer **mouse** is broken, please **replace** it” and “A **mouse** made a mess in the kitchen, please **replace** utensils”

Neural network model - Cont.

- Assign an integer to each top 1000 words
- Convert Descriptions to integers
- Padded sequences with zeros to make Descriptions same length

```
[[1]]
[1] 26 63 14 3 72 152 44 38

[[2]]
[1] 83 491 37 662 898 14 19 574 462 879 3 37 2 20 33 52 462 412 37 20 69 341

[[3]]
[1] 922 108 239 83 491 37 662 898 14 19 574 462 879 3 37 2 20 33 52 462 412 37 20 69 341

[[4]]
[1] 663 542 32 943 437 649 27 839 131 397 495 157 141 16 3 74 839 383 7 161 177 75 543 76 189 45 175 89 649
[30] 943 437 302 799 443 302 119

[[5]]
[1] 26 63 14 3 72 44 38

[[6]]
[1] 188 324
```

Neural network model - Cont.

Training the model:

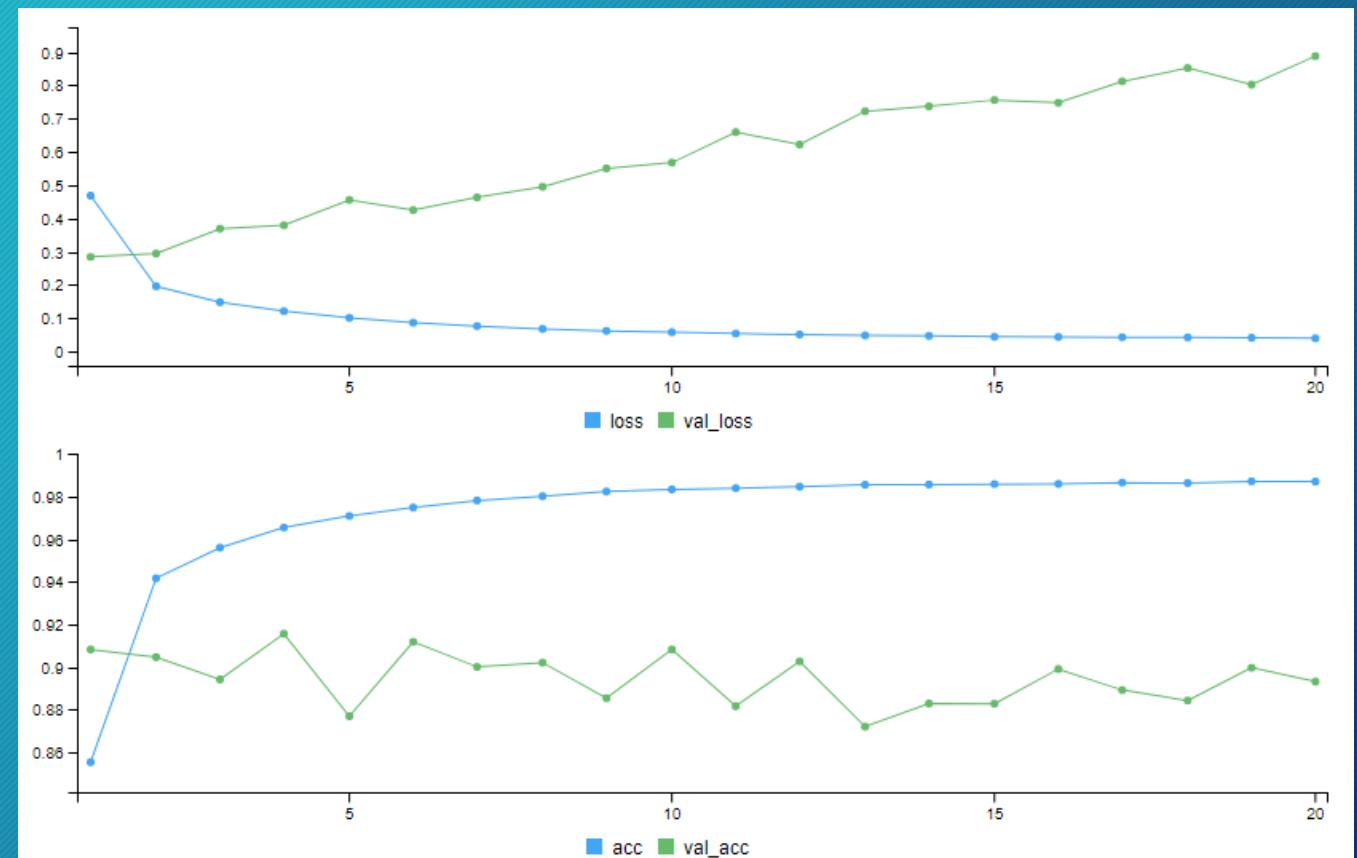
- Accuracy 97%

Overall statistics

Accuracy : 0.9707
95% CI : (0.9692, 0.9721)
No Information Rate : 0.2522
P-value [Acc > NIR] : < 2.2e-16

Kappa : 0.9657

Mcnemar's Test P-Value : NA



Neural network model - Cont.

Testing the model:

- Accuracy 90%

```
Overall statistics
  Accuracy : 0.9098
  95% CI   : (0.9047, 0.9147)
  No Information Rate : 0.2503
  P-value [Acc > NIR] : < 2.2e-16
  Kappa    : 0.8942
```

Results and conclusion

- Neural network accuracy - **90%**
- Neural network classified our target data well.

Future work

14

- Try other approaches such as Support Vector Machines.
- Research and apply advanced natural language processing techniques.
- Apply text-embeddings process to compute useful statistics such as word proximity and synonym distinction.