Our Project

SDWORX Hackathon 2021

Our Amazing Team

Kevin Albert:
Data Scientist

Hanna Oulad :
Web developer

Bayar Mohamed :
Microsoft .Net
Consultant

What is it About ?

collecting and analysing employee feedback

What is it About ?

collecting and analysing employee feedback

with the goal to explain root-causes for employee engagement

What is it About ?

collecting and analysing employee feedback

with the goal to explain root-causes for employee engagement

technical exercise :
chatbot :
cognitive services :
graph :
machine learning :

What is it About ?

technical exercise : explore the possible ?
chatbot : generate survey & collect data
cognitive services : enrich data
graph : explore data relationships
machine learning : explainability

What is it About ?

collecting and analysing employee feedback

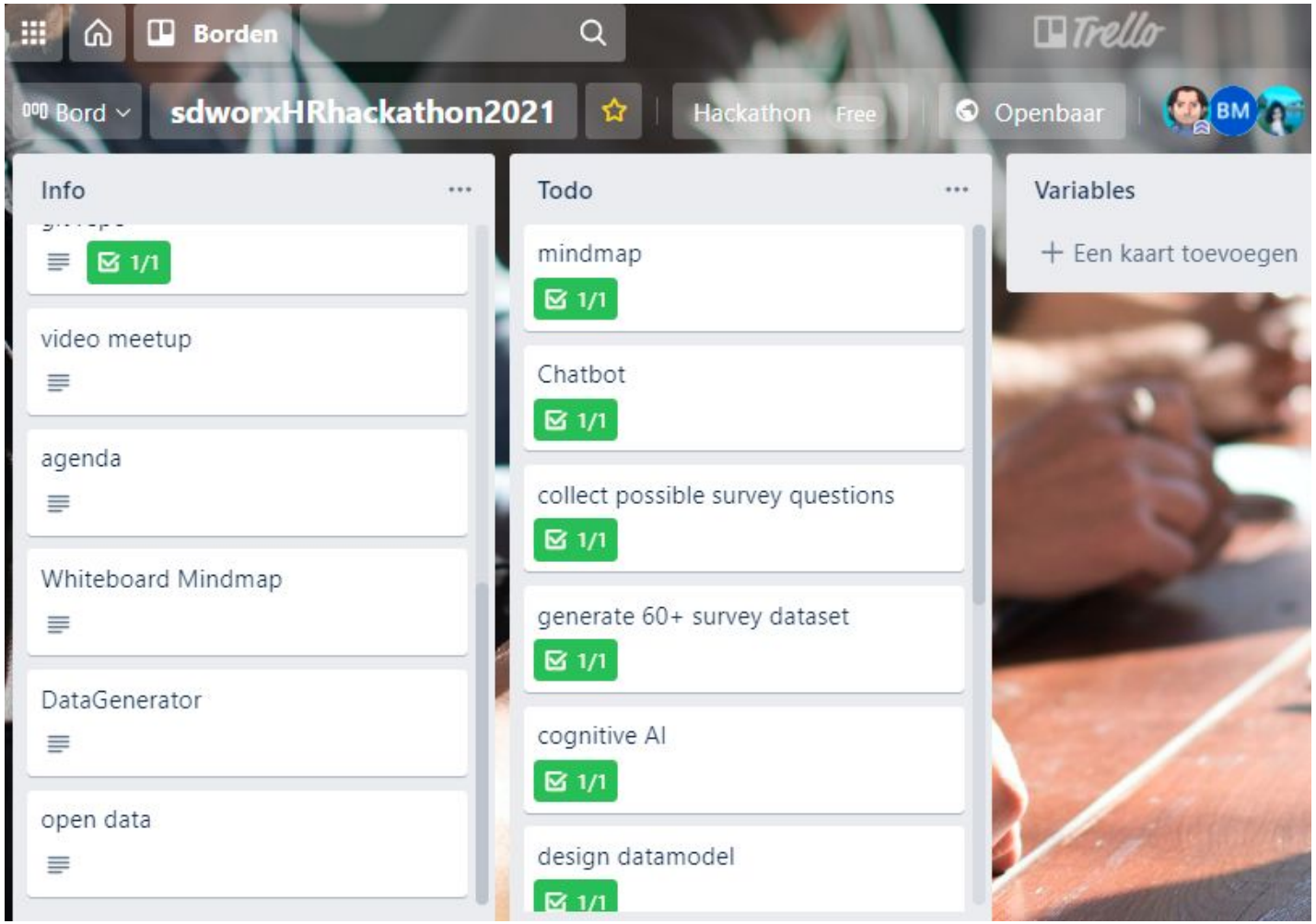
with the goal to explain root-causes for employee engagement

technical exercise : explore the possible ?
chatbot : generate survey & collect data
cognitive services : enrich data
graph : explore data relationships
machine learning : explainability

explain where a certain sentiment comes from ?
reasons for negative survey answers ?

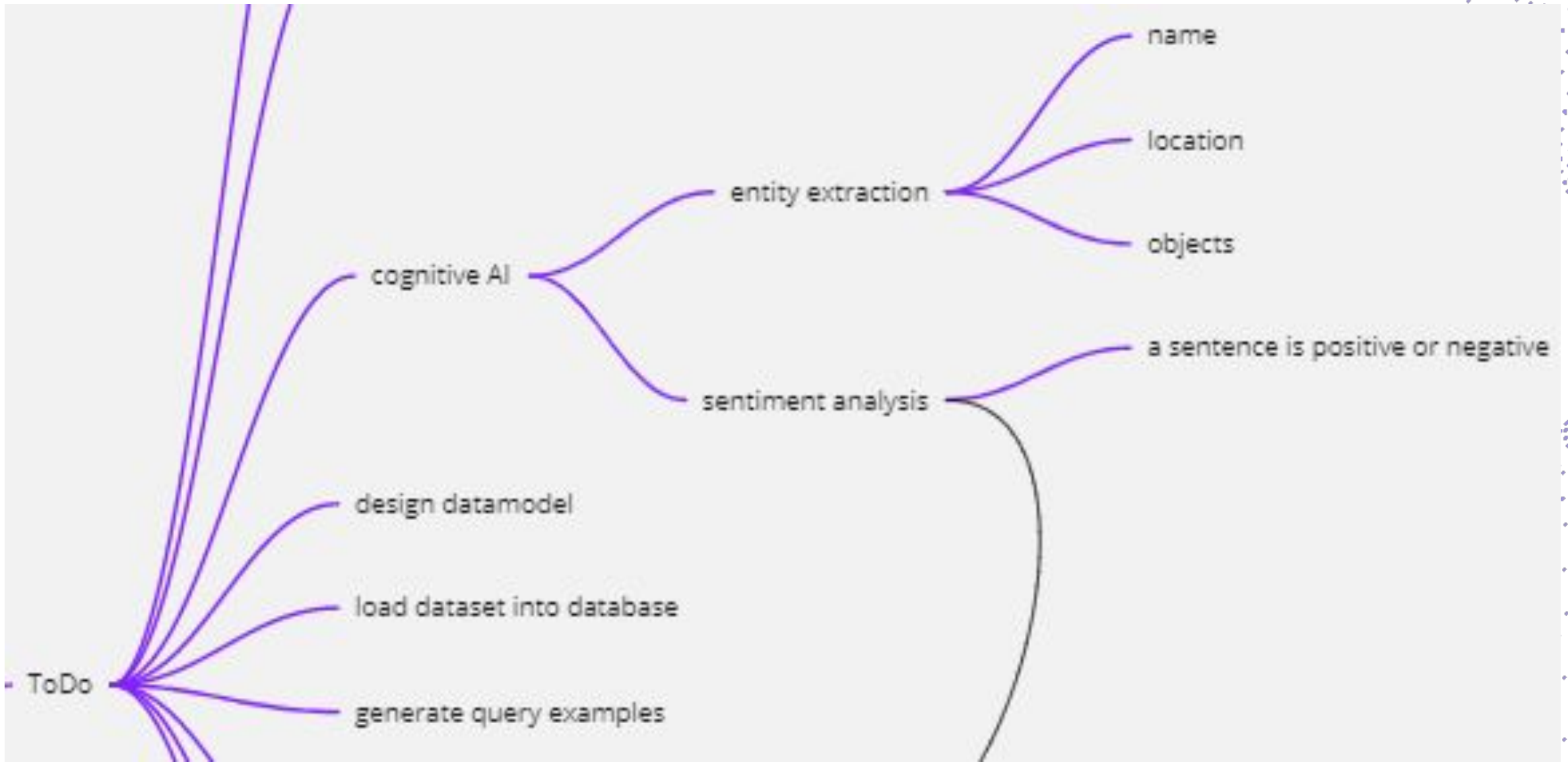
Tools

Status page



https://trello.com/b/1V3NokwO

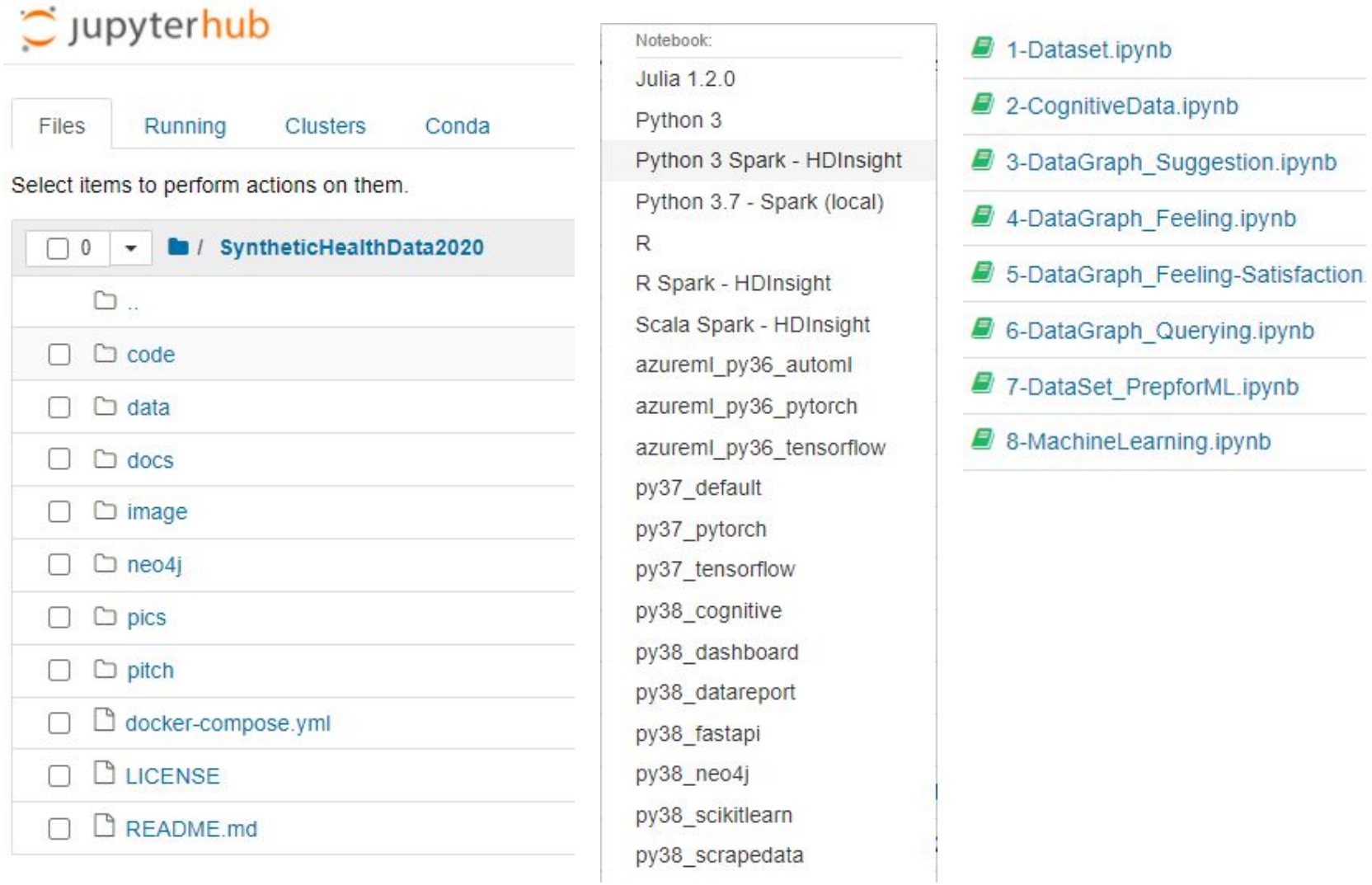
Mindmap


ToDo
cognitive AI
entity extraction
name
location
objects
sentiment analysis
a sentence is positive or negative
design datamodel
load dataset into database
generate query examples

Cloud

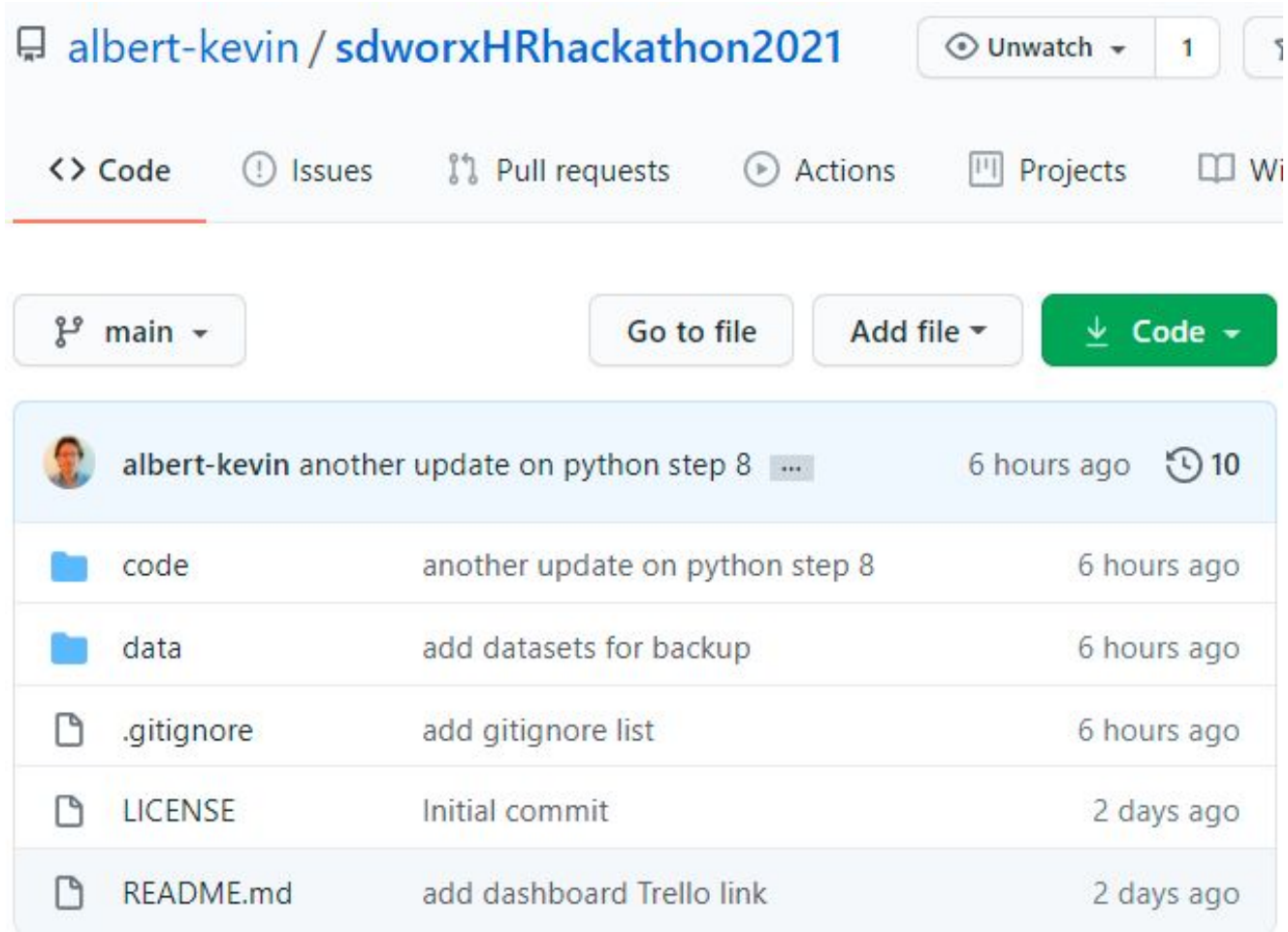
- anomalydetector27112020
- computervision27112020
- contentmoderator27112020
- customvisionPrediction27112020
- customvisionTraining27112020
- datalake27112020
- face27112020
- formrecognizer27112020
- functionApp27112020
- functionApp27112020
- languageunderstandingAuthoring27112020
- languageunderstandingPrediction27112020
- myDatabricks02
- myVM02
- myVM02_OsDisk_1_ba2
- myVM02NSG
- myVM02PublicIP
- myVM02VMNic
- machine_learning_workspace02
- machinelinsights73d0a6bb
- machinelkeyvaulte374e73a
- machinelstorage9af0d08f1
- myADF27112020

Development environment



Git repo

https://github.com/albert-kevin/sdworxHRhackathon2021
https://github.com/bayarmohamed/Hackathon



Synthetic survey data

How was your last month in your assignment?

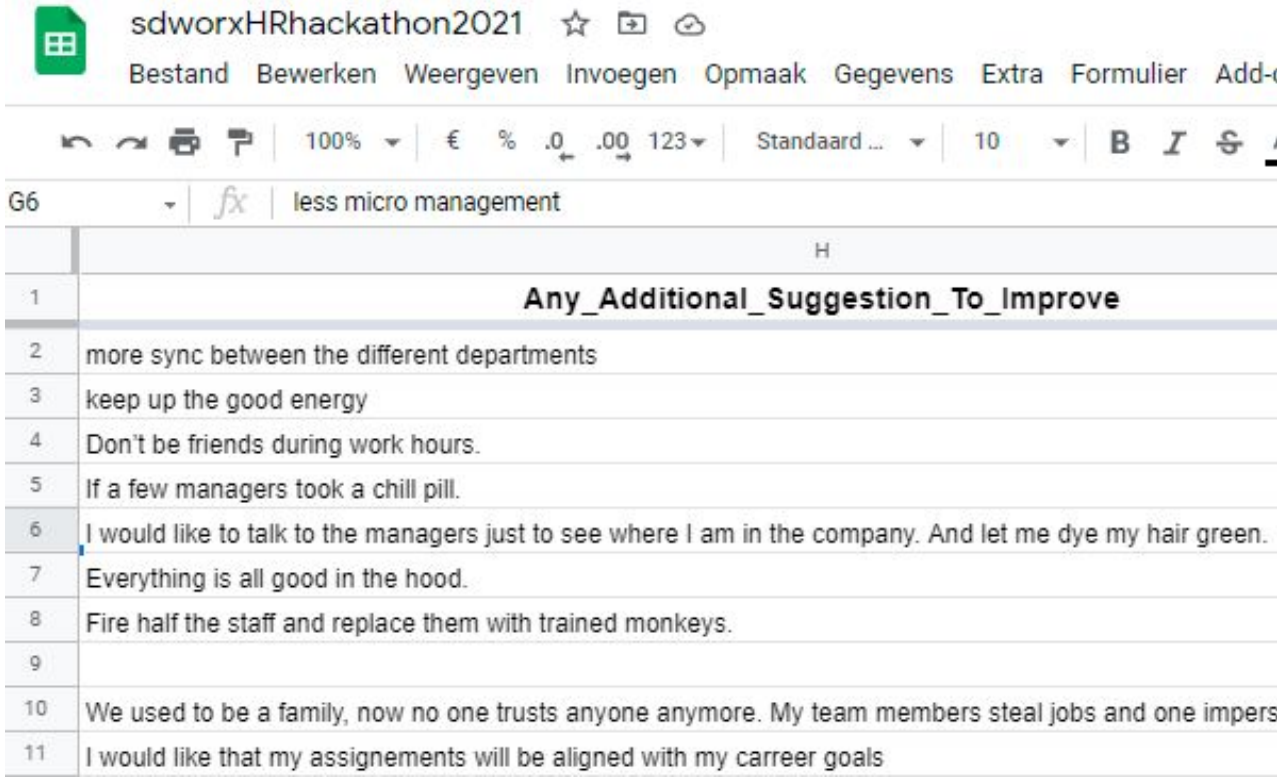
	1	2	3	4	
Not Happy at all 😖	◯	◯	◯	◯	Very Happy 😁

How was your last month within your Pod ?

	1	2	3	4	
Not Happy at all 😖	◯	◯	◯	◯	Very Happy 😁

Synthetic survey data

```
Data columns (total 11 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   email                                   100 non-null    object
 1   full_name                               100 non-null    object
 2   how_was_your_last_month_in_your_assignment  100 non-null    int64
 3   How_Was_Your_Last_Month_Within_Your_Department  100 non-null    int64
 4   How_Was_Your_Last_Month_With_Us         100 non-null    int64
 5   What_Felt_Best_During_This_Last_Month   37 non-null     object
 6   What_Should_We_Do_To_Make_You_Feel_Better_Satisfaction  37 non-null
 7   Any_Additional_Suggestion_To_Improve    15 non-null
 8   any_additional_comments                 0 non-null
 9   I_Would_Like_To_Get_Called_By           0 non-null
 10  I_Would_Like_To_Provide_A_Copy_Of_My_Answers_Tc  0 non-null
dtypes: float64(3), int64(3), object(5)
memory usage: 8.7+ KB
```



	H
1	**Any_Additional_Suggestion_To_Improve**
2	more sync between the different departments
3	keep up the good energy
4	Don't be friends during work hours.
5	If a few managers took a chill pill.
6	I would like to talk to the managers just to see where I am in the company. And let me dye my hair green.
7	Everything is all good in the hood.
8	Fire half the staff and replace them with trained monkeys.
9	
10	We used to be a family, now no one trusts anyone anymore. My team members steal jobs and one impers
11	I would like that my assignements will be aligned with my carreer goals

Entity extraction

	document	entity_text	entity_category
0	keep the good vibes, the trust and open commun...	communication	Skill
1	maintain the good collaboration	collaboration	Skill
2	more consideration for the employes, making th...	management decision	Skill
3	less micro management	micro management	Skill
4	valuate all staff the same way	valuate	Skill
5	more transparency about the financial situatio...	financial	Skill

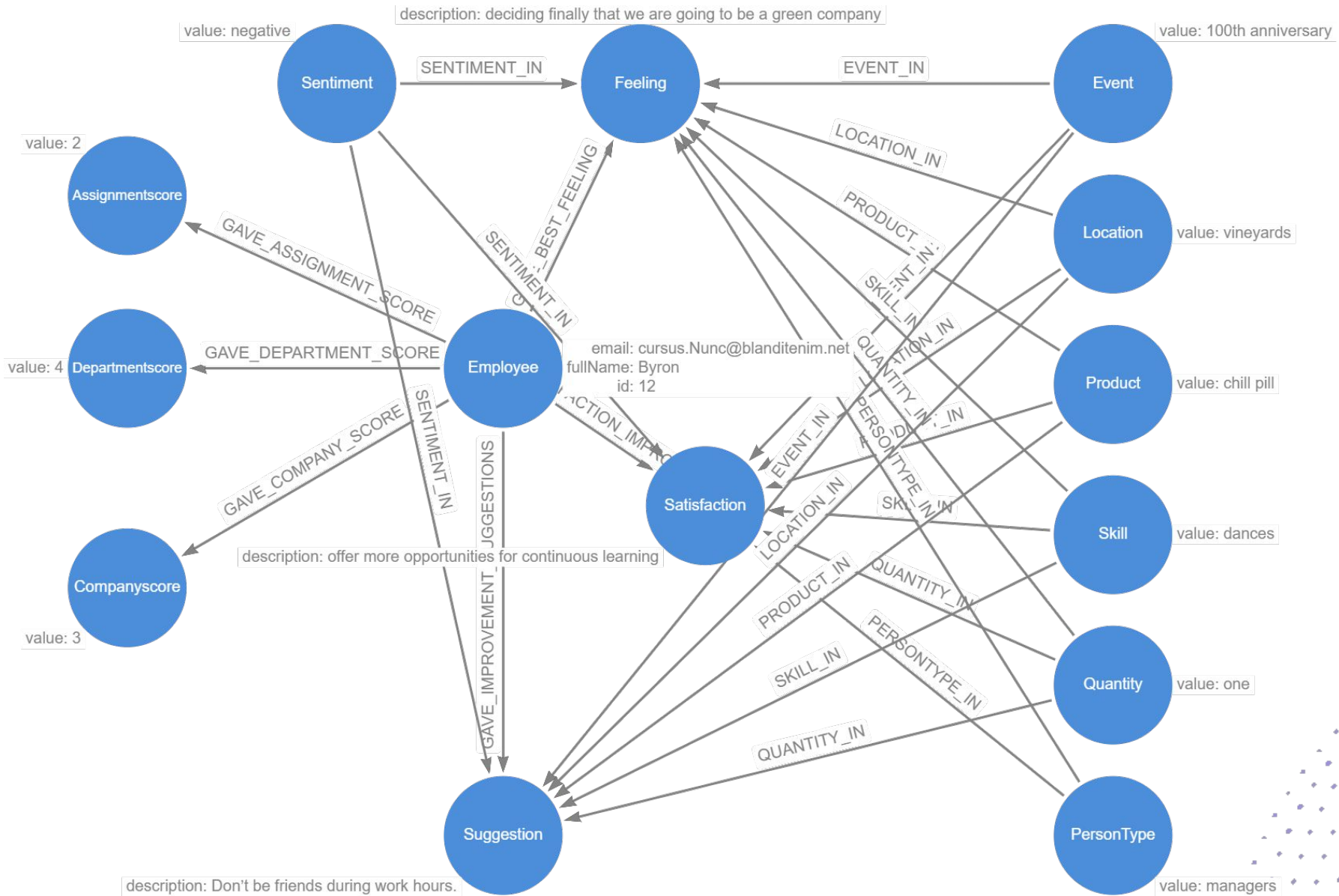
Sentiment extraction

	document	sentiment
0	keep the good vibes, the trust and open commun...	positive
1	maintain the good collaboration	positive
2	more consideration for the employes, making th...	neutral
3	I don't feel that my skills and capacities are...	negative
4	less micro management	neutral

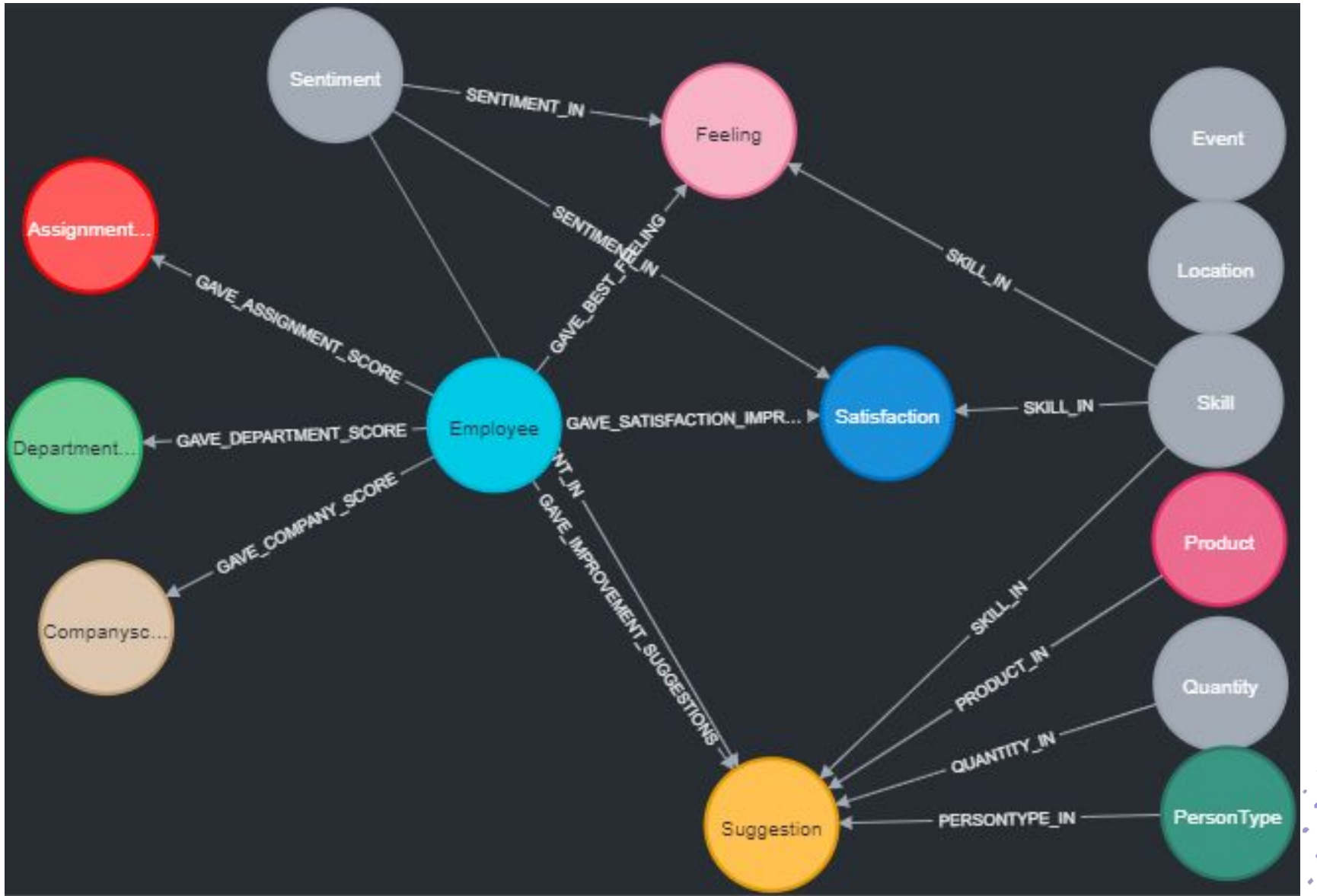
Key phrases extraction

	document	key_phrases
0	keep the good vibes, the trust and open commun...	['trust', 'good vibes', 'open communication']
1	maintain the good collaboration	['good collaboration']
2	more consideration for the employes, making th...	['employes', 'partners', 'management decision'...
3	I don't feel that my skills and capacities are...	['skills', 'capacities', 'accounts']
4	less micro management	['micro management']

Datamodel


description: deciding finally that we are going to be a green company
value: negative
Sentiment
SENTIMENT_IN
Feeling
EVENT_IN
Event
value: 100th anniversary
value: 2
Assignmentscore
GAVE_ASSIGNMENT_SCORE
LOCATION_IN
PRODUCT_IN
Location
value: vineyards
SENTIMENT_IN
BEST_FEELING
SKILL_IN
value: 4
Departmentscore
GAVE_DEPARTMENT_SCORE
Employee
email: cursus.Nunc@blanditenim.net
fullName: Byron
id: 12
Product
value: chill pill
GAVE_COMPANY_SCORE
SENTIMENT_IN
Satisfaction
EVENT_IN
LOCATION_IN
Skill
value: dances
description: offer more opportunities for continuous learning
Companyscore
value: 3
GAVE_IMPROVEMENT_SUGGESTIONS
PRODUCT_IN
QUANTITY_IN
PERSONTYPE_IN
SKILL_IN
Quantity
value: one
QUANTITY_IN
Suggestion
PersonType
description: Don't be friends during work hours.
value: managers

Graph database



Some Cool Stuff We Did

what percentages has a positive, negative and neutral sentiment ?

total sentiment percentage ?

```
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->()<-[rel]-(e:Employee)
WITH count(s) AS totalPositive
MATCH (s:Sentiment {value:'negative'})-[:SENTIMENT_IN]->()<-[rel]-(e:Employee)
WITH totalPositive, count(s) AS totalNegative
MATCH (s:Sentiment {value:'neutral'})-[:SENTIMENT_IN]->()<-[rel]-(e:Employee)
WITH totalPositive, totalNegative, count(s) AS totalNeutral
MATCH (s:Sentiment)-[:SENTIMENT_IN]->()<-[rel]-(e:Employee)
WITH totalPositive, totalNegative, totalNeutral, count(s) AS total
RETURN toFloat((totalPositive)/toFloat(total))*100 AS pos,
toFloat((totalNegative)/toFloat(total))*100 AS neg,
toFloat((totalNeutral)/toFloat(total))*100 AS neu
```

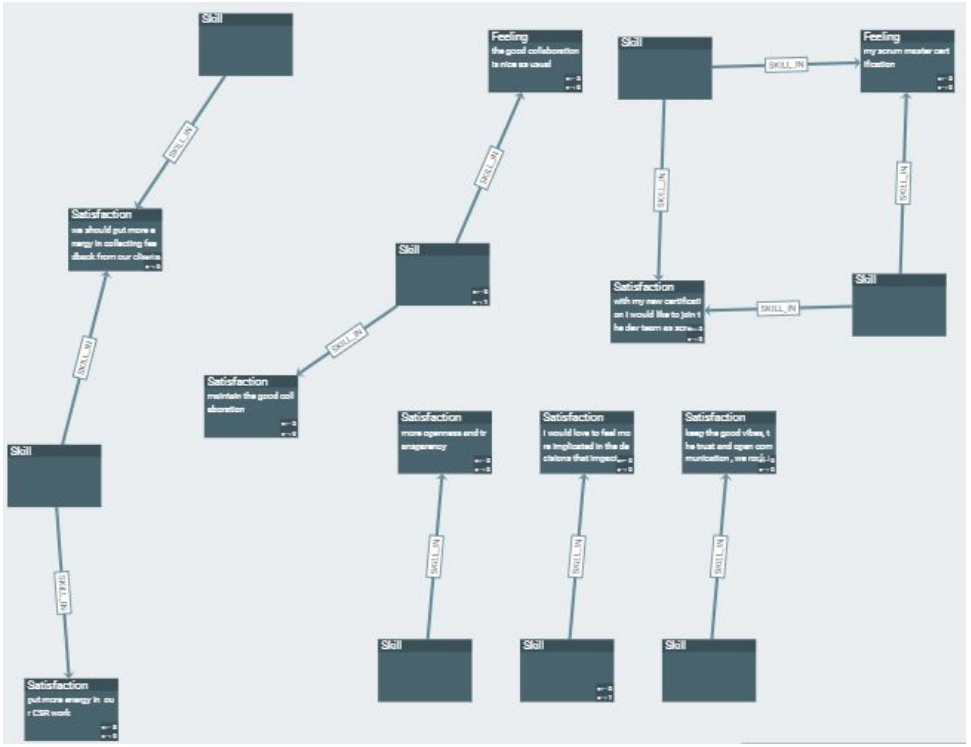
	pos	neg	neu
0	49.438202	13.483146	37.078652

what percentages has a positive, negative and neutral sentiment ?

positive sentiment entity words ?

	positive_entities
0	collecting
1	energy
2	openness
3	impact
4	certification
5	scrum
6	CSR
7	collaboration
8	communication

```
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->(feeling:Feeling)
WITH feeling
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->(satisfaction:Satisfaction)
WITH feeling, satisfaction
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->(suggestion:Suggestion)
WITH feeling, satisfaction, suggestion
MATCH (satisfaction)<-[r:SKILL_IN]-(d)
RETURN DISTINCT d.value AS positive_entities
```



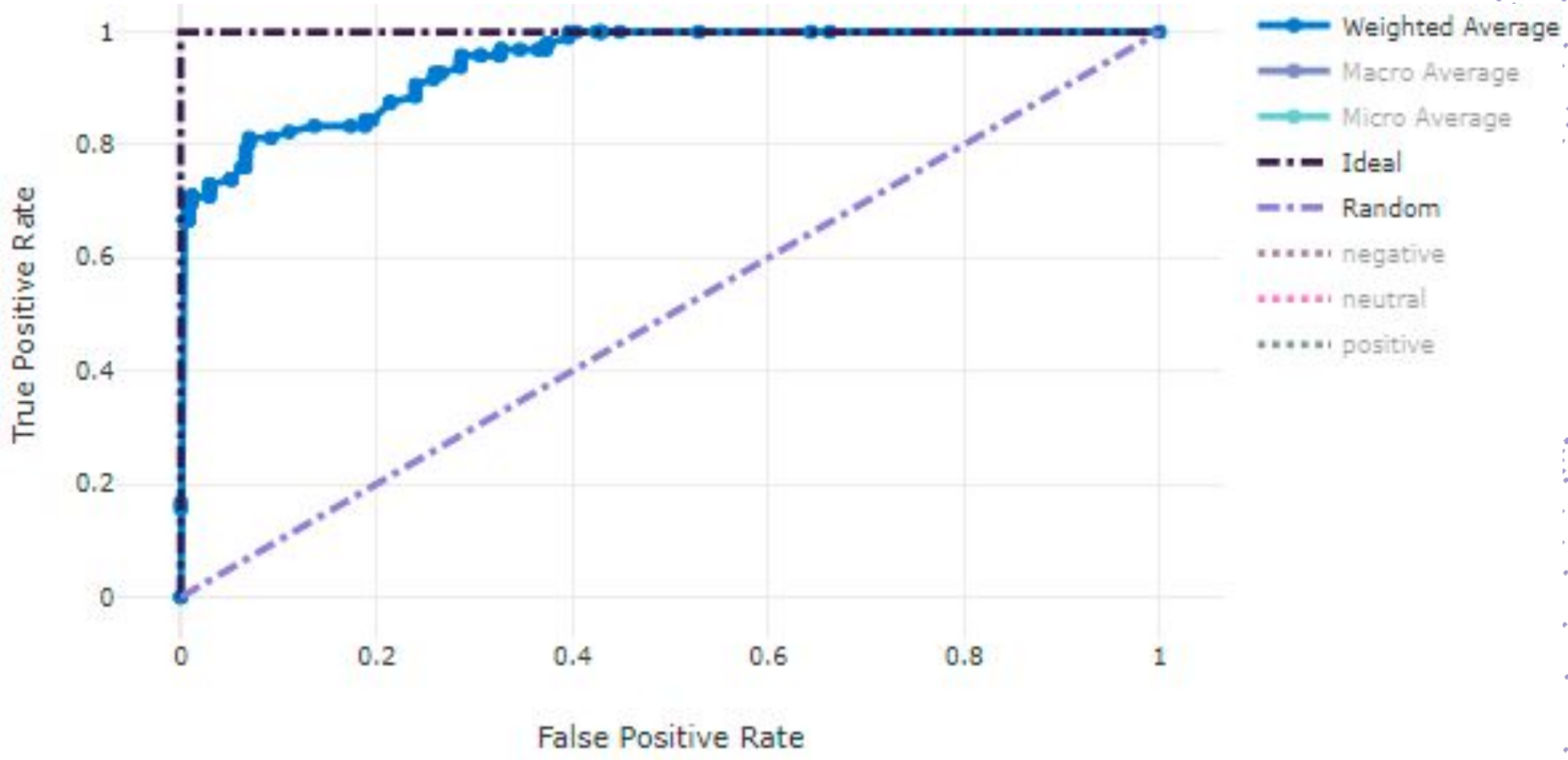
what percentages has a positive, negative and neutral sentiment ?

cluster ?

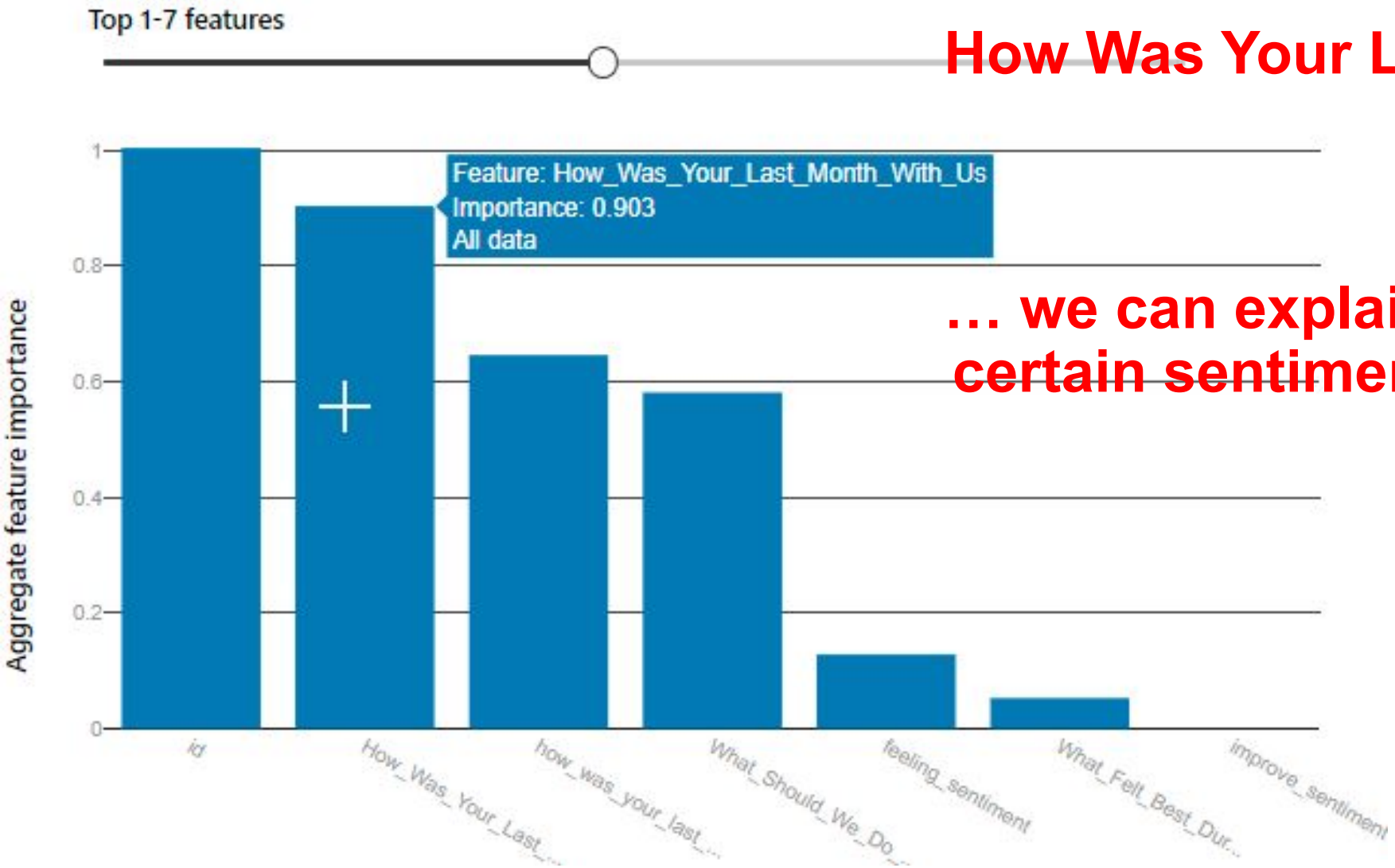
Machine Learning

ITERATION	PIPELINE	DURATION	METRIC	BEST
0	MaxAbsScaler LightGBM	0:00:19	0.9312	0.9312
1	MaxAbsScaler XGBoostClassifier	0:00:18	0.9450	0.9450
2	MaxAbsScaler RandomForest	0:00:19	0.9144	0.9450
3	MaxAbsScaler RandomForest	0:00:29	0.9213	0.9450
4	MaxAbsScaler RandomForest	0:00:19	0.9419	0.9450
5	MaxAbsScaler ExtremeRandomTrees	0:00:19	0.9345	0.9450
6	MaxAbsScaler ExtremeRandomTrees	0:00:20	0.9177	0.9450
7	MaxAbsScaler ExtremeRandomTrees	0:00:19	0.9098	0.9450
8	VotingEnsemble	0:02:32	0.9651	0.9651
9	StackEnsemble	0:02:32	0.9713	0.9713

Machine Learning



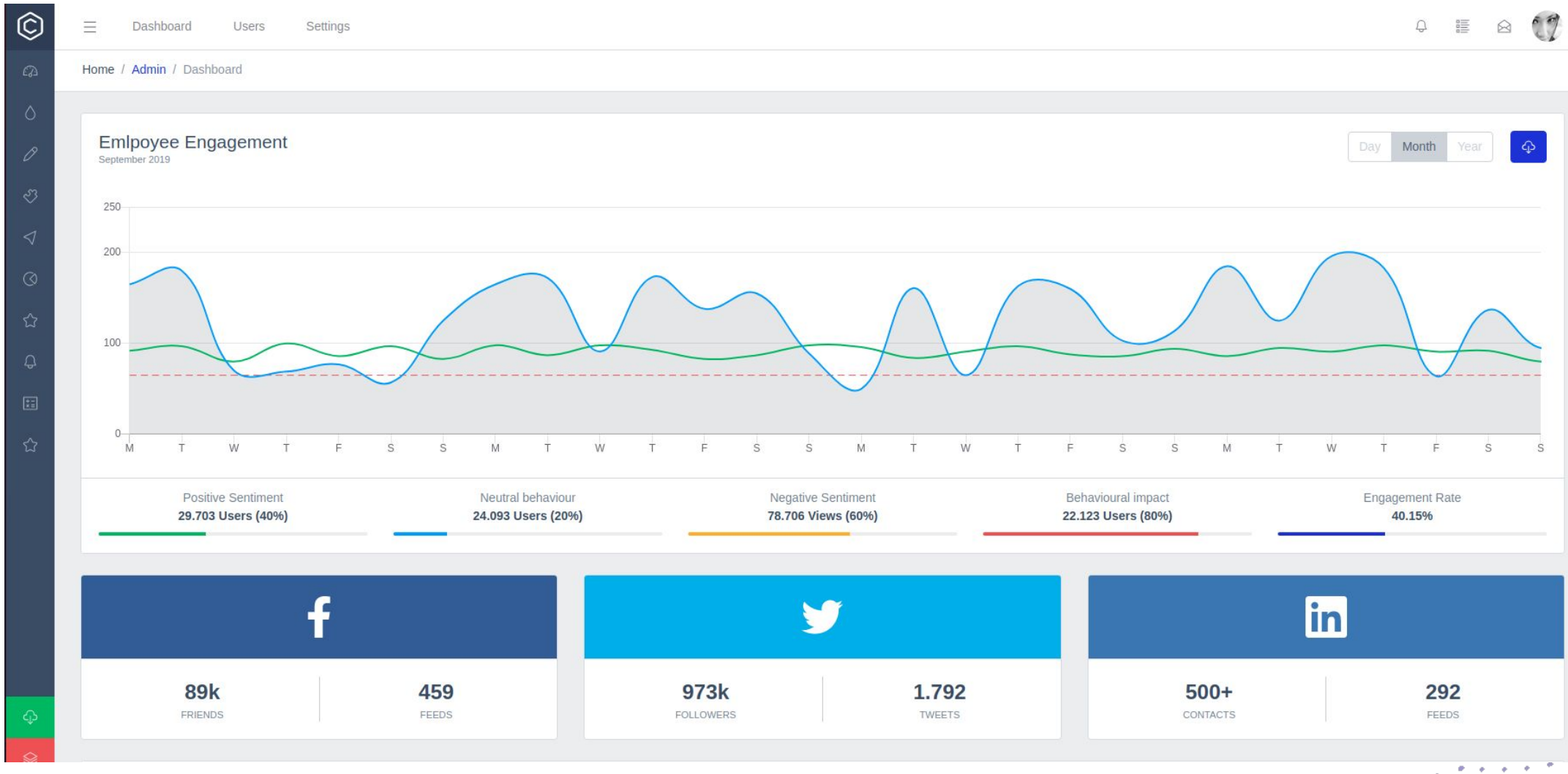
Most important Sentiment Influencer



How Was Your Last Month With Us ?

… we can explain why a cluster or a certain sentiment comes from

Showcasing the data treatment



Requirements

Asp.net core web api project with

```
Microsoft.Bot.Builder.Integration.AspNet.Core
```

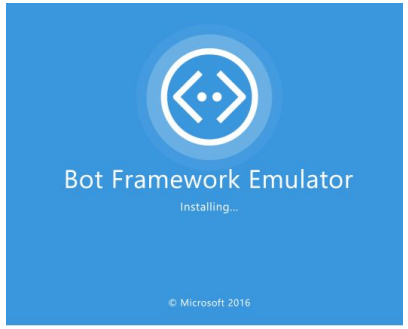
Bot framework Emulator

Entity Framework Core

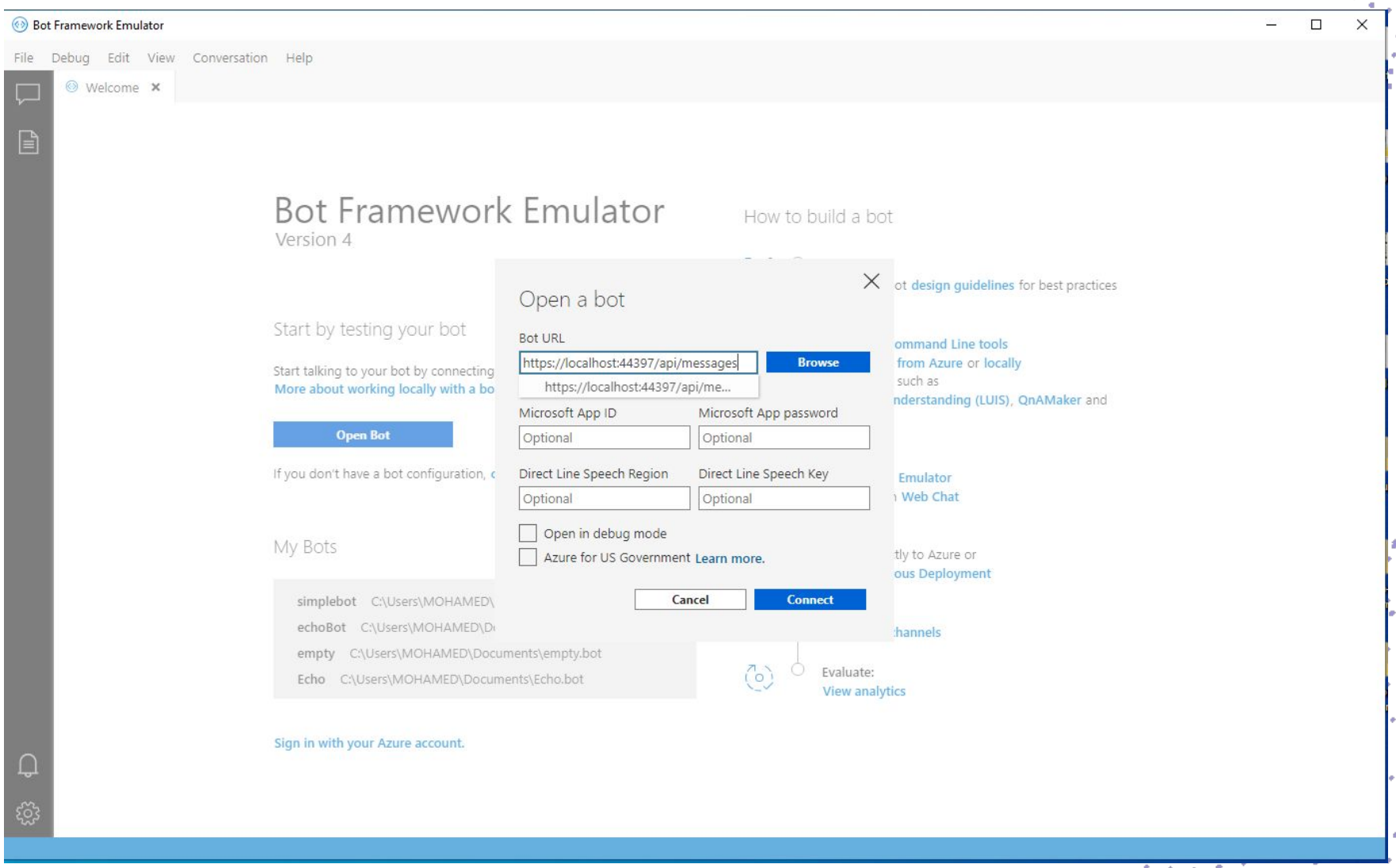
Sql Server

adaptivecards.io web site

source code : [bayarmohamed/Hackathon (github.com)]

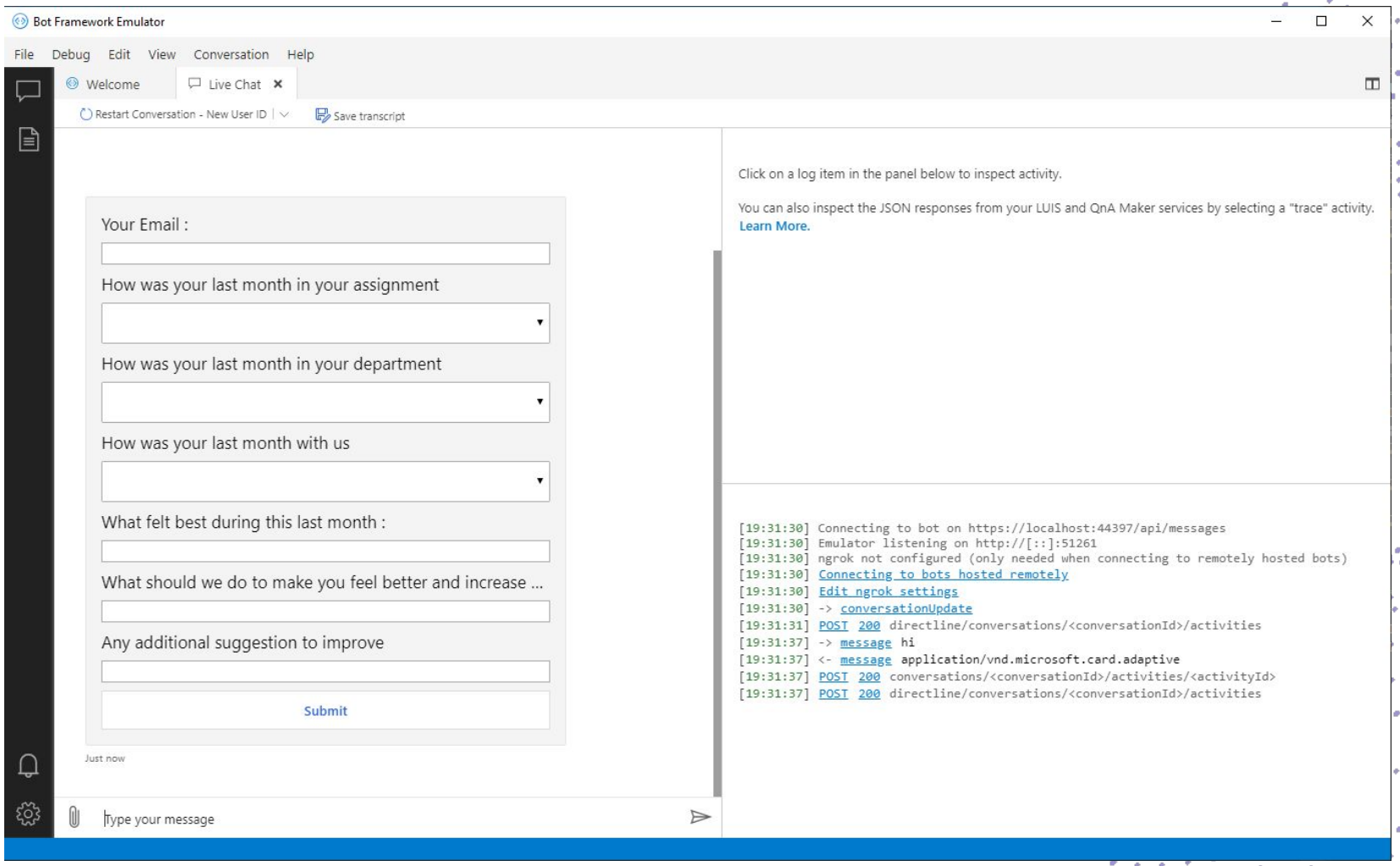


Bot Framework Emulator



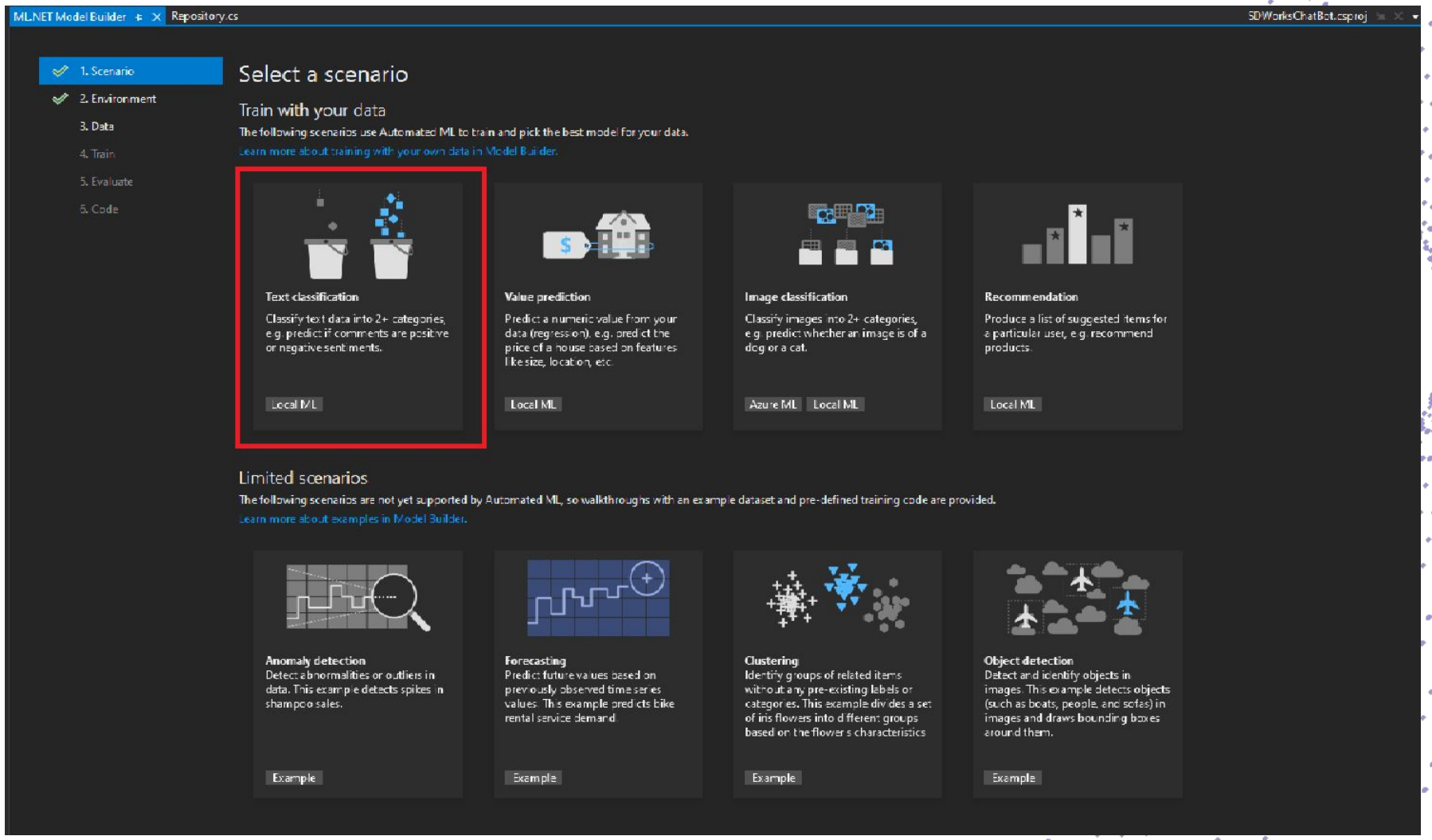
ChatBot For Survey

Use chatbot with “Adaptive Cards” SDK or json to generate a survey form

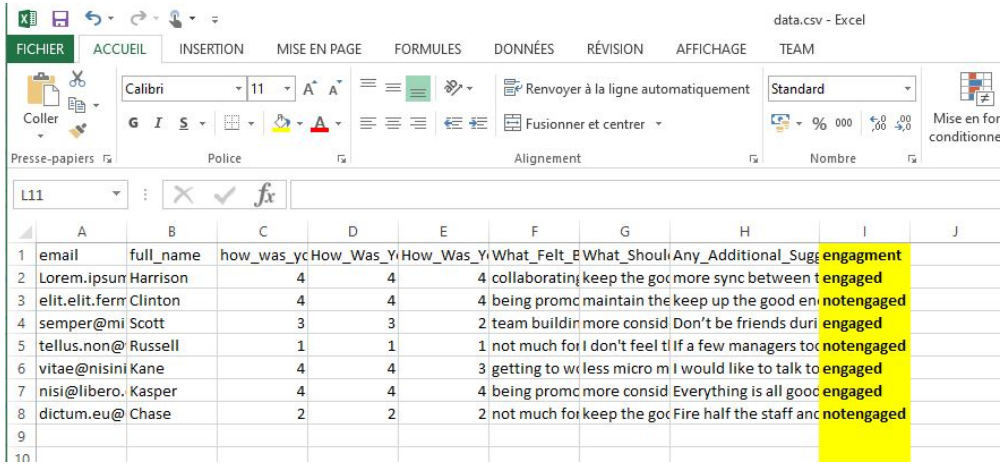


Training Data with ML.Net

Select a binary classification as a scenario

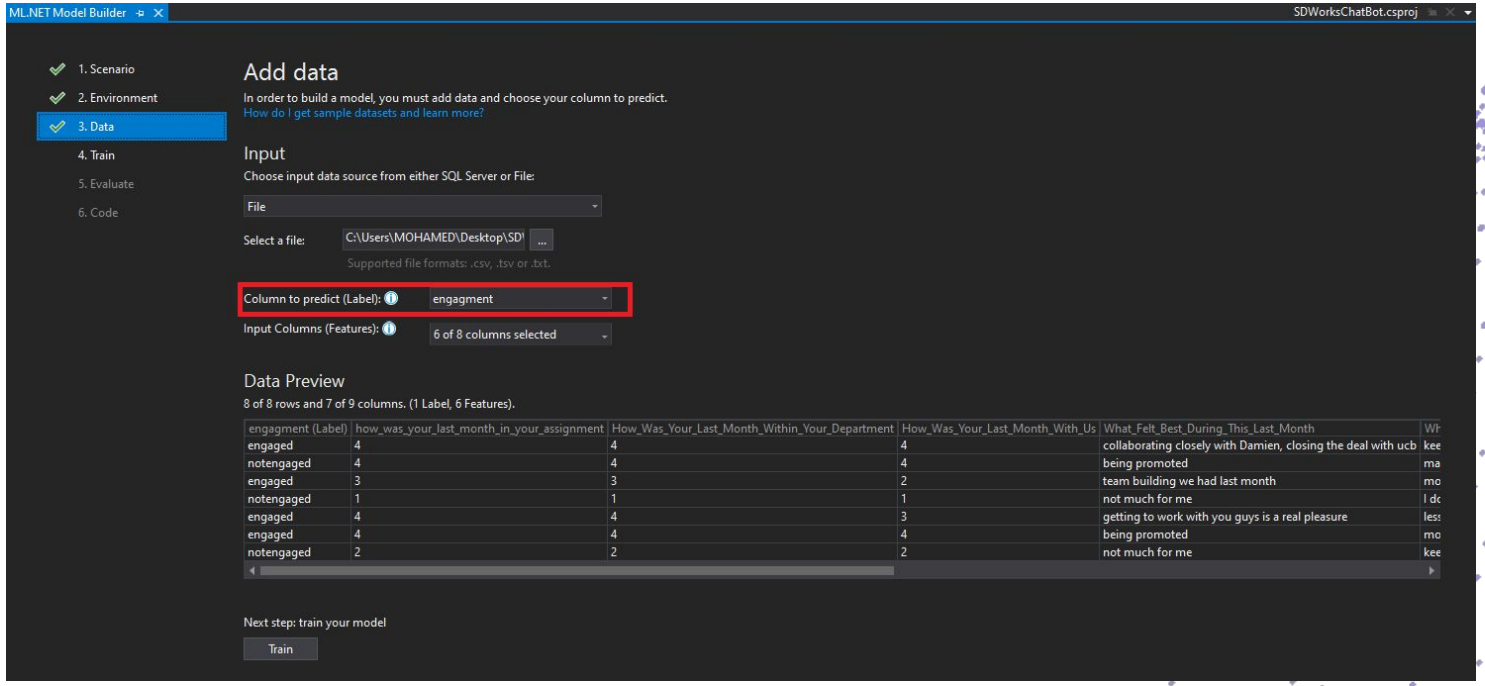


Training Data



	A	B	C	D	E	F	G	H	I	J
1	email	full_name	how_was_yc	How_Was_Y	How_Was_Y	What_Felt_E	What_Shoul	Any_Additional_Sugg	engagment	
2	Lorem.ipsun	Harrison	4	4	4	collaboratin	keep the goc	more sync between t	engaged	
3	elit.elit.ferm	Clinton	4	4	4	being promc	maintain the	keep up the good en	notengaged	
4	semper@mi	Scott	3	3	2	team buildin	more consid	Don't be friends duri	engaged	
5	tellus.non@	Russell	1	1	1	not much for	I don't feel tl	If a few managers toc	notengaged	
6	vitae@nisini	Kane	4	4	3	getting to wc	less micro m	I would like to talk to	engaged	
7	nisi@libero.	Kasper	4	4	4	being promc	more consid	Everything is all gooc	engaged	
8	dictum.eu@	Chase	2	2	2	not much for	keep the goc	Fire half the staff anc	notengaged	
9										
10										

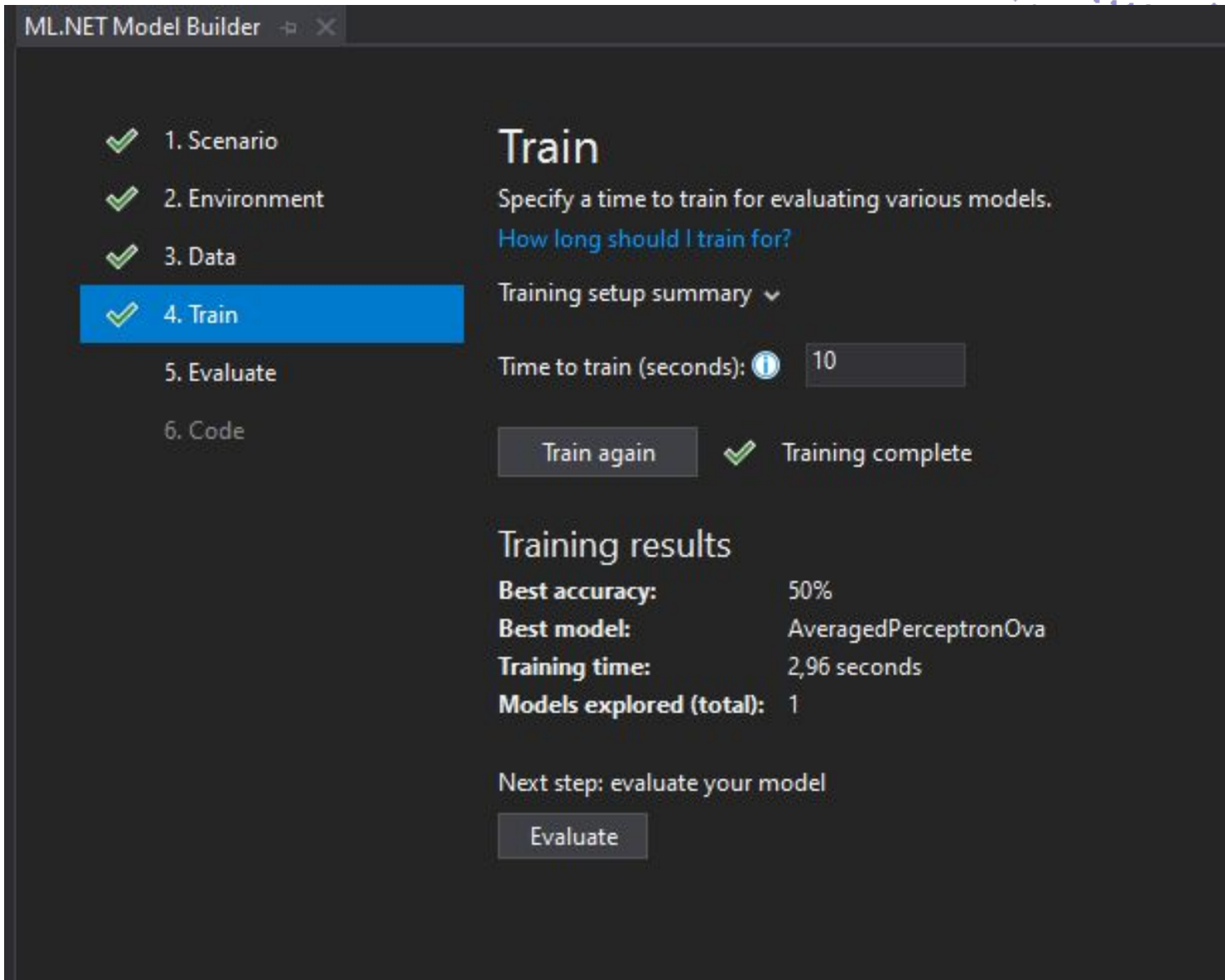
Import Data from csv file
engagement field is a Label
Features are other fields except email and name



Train the model

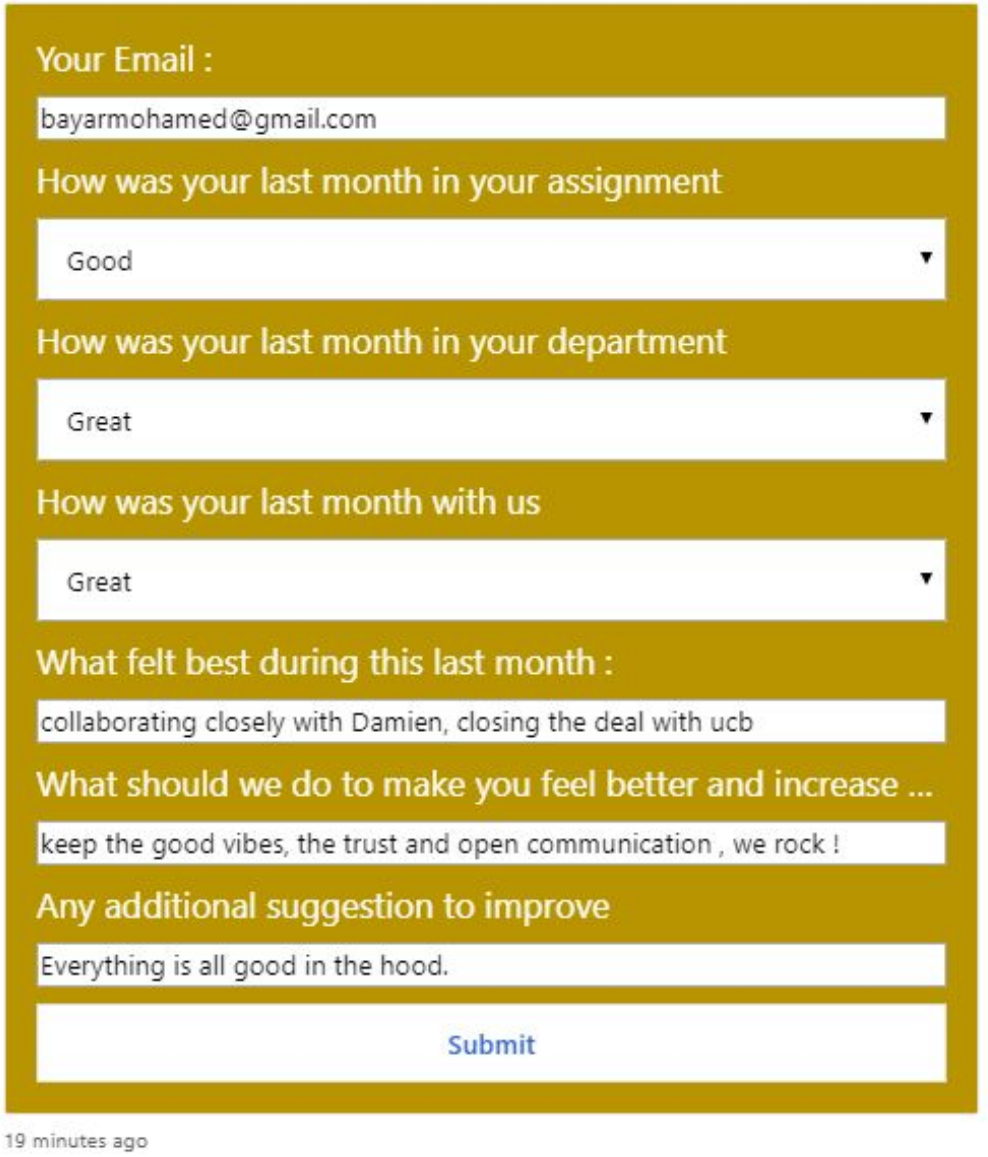
Train : generate a model to be used to predict survey
The training select the more adapted model (best model)
In our case the **AveragedPerceptronOva.**

The accuracy : 50% (Accuracy is the proportion of correct predictions with a test data set. It is the ratio of number of correct predictions to the total number of input samples. It works well if there are similar number of samples belonging to each class)



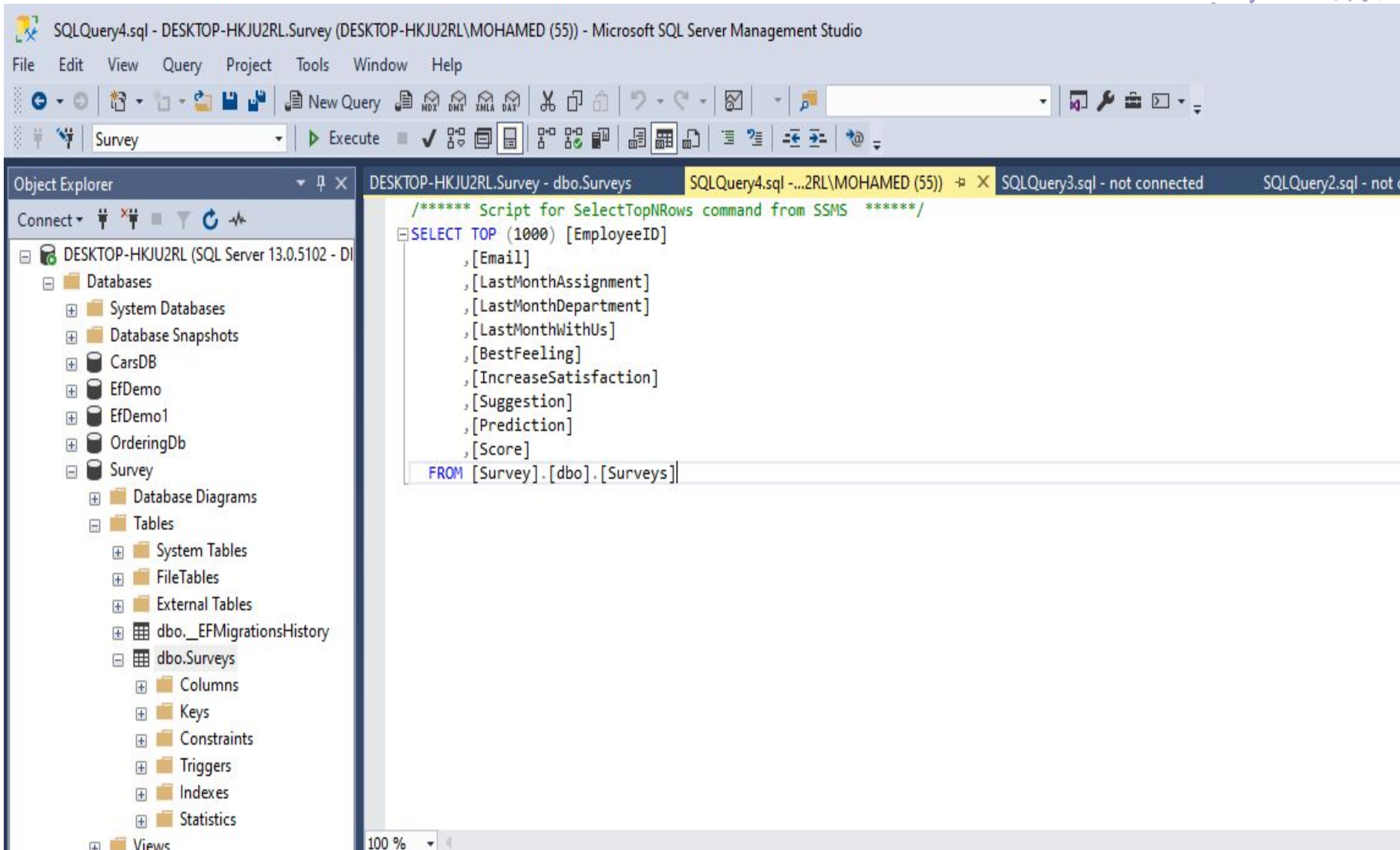
Predict the survey

F1 score also known as *balanced F-score or F-measure*. It's the harmonic mean of the precision and recall. F1 Score is helpful when you want to seek a balance between Precision and Recall.
The closer to 1.00, the better. An F1 score reaches its best value at 1.00 and worst score at 0.00. It tells you how precise your classifier is.



Save predicted surveys

sql server database is used to save predicted surveys



Questions

Kevin Albert:
Data Scientist

Hanna Oulad:
Web developer

Bayar Mohamed:
MS .Net Consultant