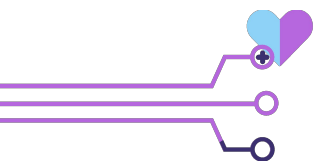


Our Project

SDWORX Hackathon 2021





Our Amazing Team

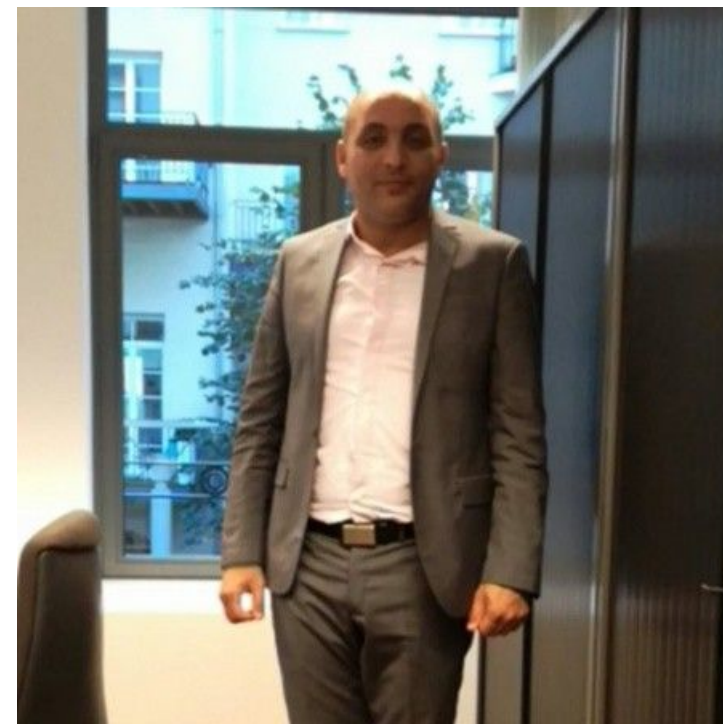
Kevin Albert:
Data Scientist

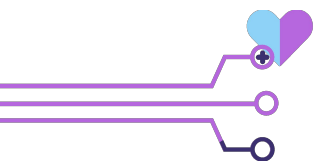


Hanna Oulad :
Web developer



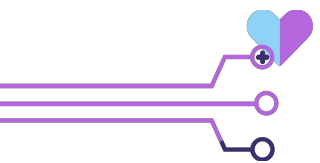
Bayar Mohamed :
Microsoft .Net
Consultant





What is it About ?

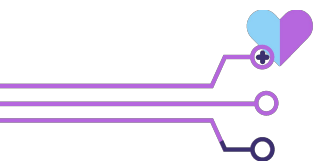
collecting and analysing employee feedback



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

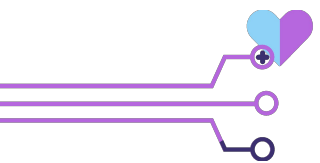
technical exercise :

chatbot :

cognitive services :

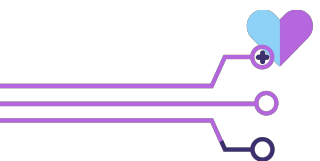
graph :

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

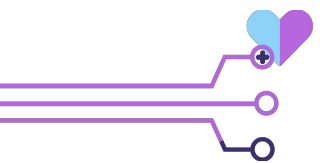
cognitive services : enrich data

graph : explore data relationships

machine learning : explainability

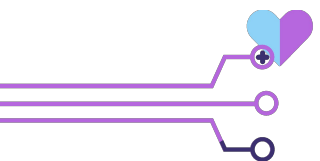
explain where a certain sentiment comes from ?

reasons for negative survey answers ?

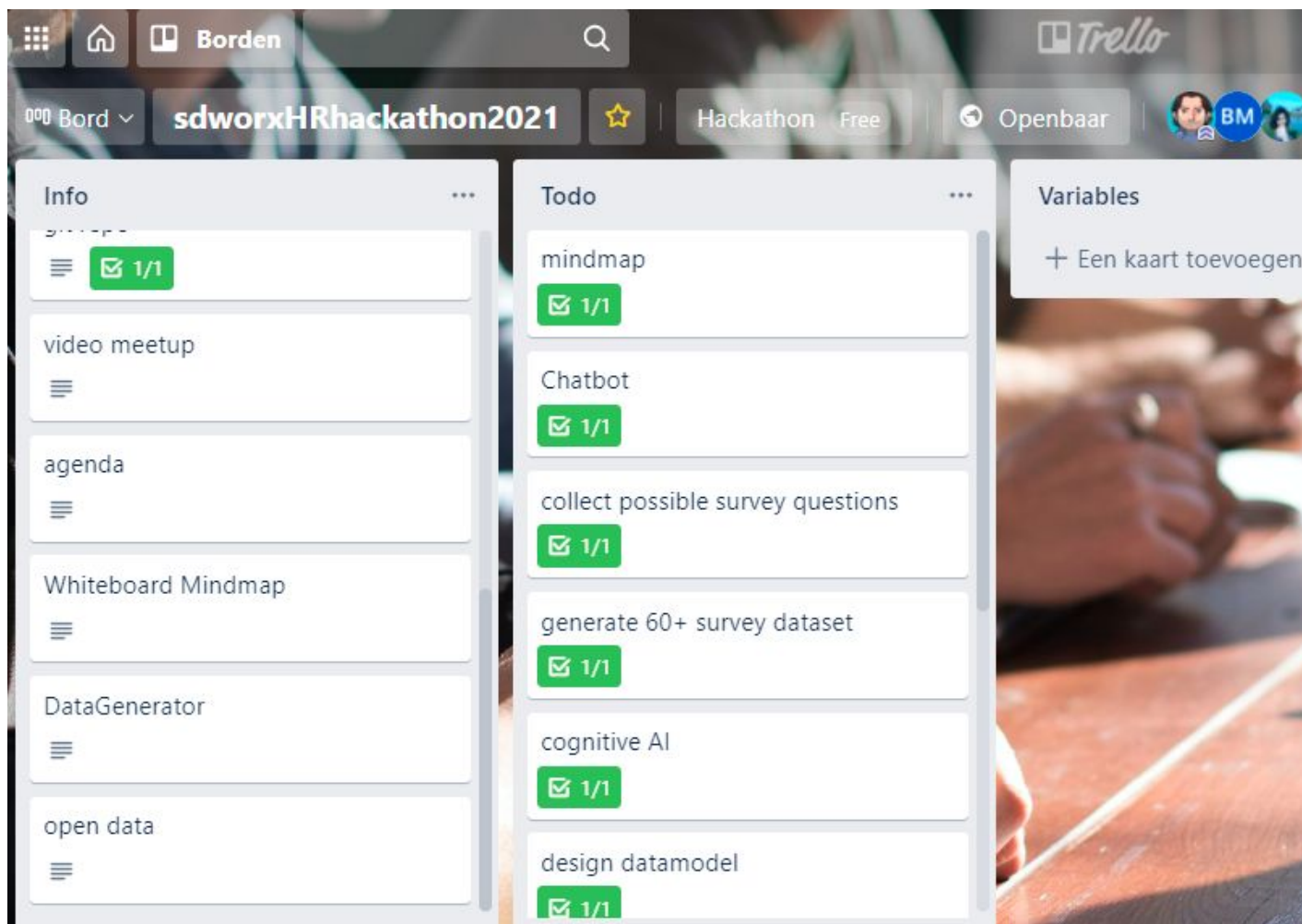


Tools

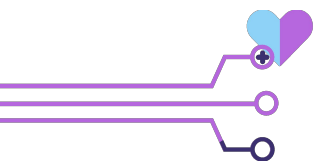




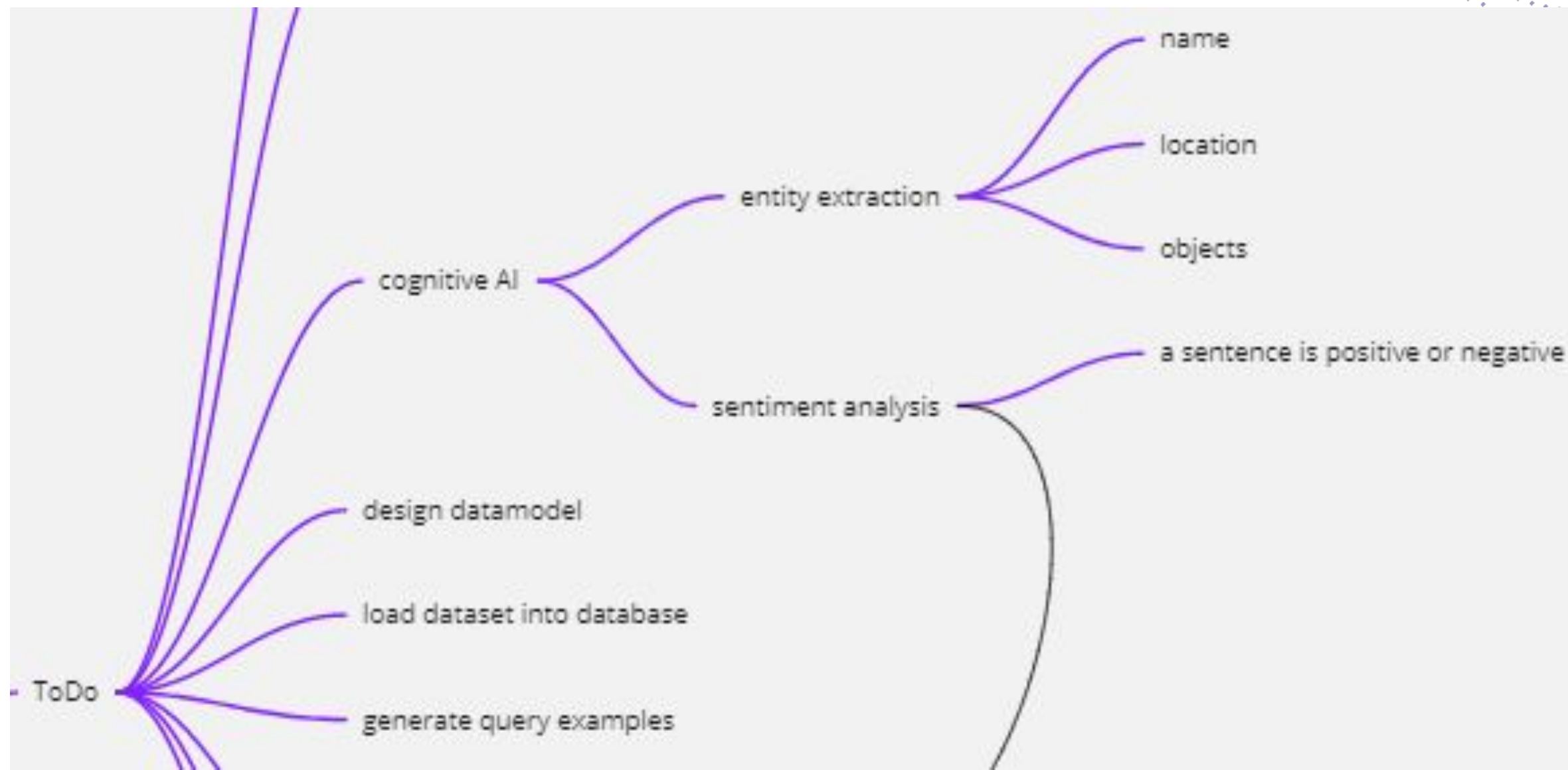
Status page

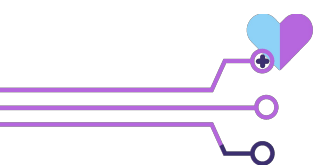


<https://trello.com/b/1V3NokwO>









Mindmap





Cloud

-  anomalydetector27112020
-  computervision27112020
-  contentmoderator27112020
-  customvisionPrediction27112020
-  customvisionTraining27112020
-  datalake27112020

 face27112020

 formrecognizer27112020

 functionApp27112020

 functionApp27112020

 languageunderstandingAuthoring27112020

 languageunderstandingPrediction27112020

 myDatabricks02

 myVM02

 myVM02_OsDisk_1_ba2

 myVM02NSG

 myVM02PublicIP

 myVM02VMNic

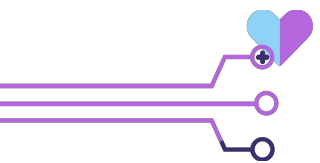
 machine_learning_workspace02

 machinelinsights73d0a6bb

 machinekeyvaulte374e73a

 machinelstorage9af0d08f1

 myADF27112020



Development environment



Files Running Clusters Conda

Select items to perform actions on them.

☐ 0 / SyntheticHealthData2020

..

☐ code

☐ data

☐ docs

☐ image

☐ neo4j

☐ pics

☐ pitch

☐ docker-compose.yml

☐ LICENSE

☐ README.md

Notebook:

Julia 1.2.0

Python 3

Python 3 Spark - HDInsight

Python 3.7 - Spark (local)

R

R Spark - HDInsight

Scala Spark - HDInsight

azureml_py36_automl

azureml_py36_pytorch

azureml_py36_tensorflow

py37_default

py37_pytorch

py37_tensorflow

py38_cognitive

py38_dashboard

py38_datareport

py38_fastapi

py38_neo4j

py38_scikitlearn

py38_scrapedata

1-Dataset.ipynb

2-CognitiveData.ipynb

3-DataGraph_Suggestion.ipynb

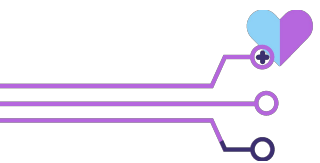
4-DataGraph_Feeling.ipynb

5-DataGraph_Feeling-Satisfaction

6-DataGraph_Querying.ipynb

7-DataSet_PrepforML.ipynb

8-MachineLearning.ipynb



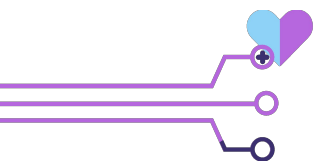
Git repo

<https://github.com/albert-kevin/sdworxHRhackathon2021>

<https://github.com/bayarmohamed/Hackathon>

The screenshot shows the GitHub interface for the repository 'albert-kevin / sdworxHRhackathon2021'. At the top, there are tabs for 'Code', 'Issues', 'Pull requests', 'Actions', 'Projects', and 'Wiki'. Below the tabs, there are buttons for 'main', 'Go to file', 'Add file', and a green 'Code' button. The main content area shows a commit history table with columns for the commit author, the commit message, and the time since the commit.

Commit	Message	Time
albert-kevin	another update on python step 8	6 hours ago
code	another update on python step 8	6 hours ago
data	add datasets for backup	6 hours ago
.gitignore	add gitignore list	6 hours ago
LICENSE	Initial commit	2 days ago
README.md	add dashboard Trello link	2 days ago



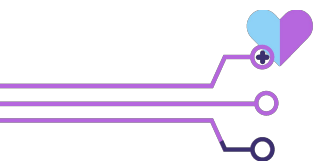
Synthetic survey data

How was your last month in your assignment?

	1	2	3	4	
Not Happy at all 😞	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Happy 😄

How was your last month within your Pod ?

	1	2	3	4	
Not Happy at all 😞	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Happy 😄



Synthetic survey data

Data columns (total 11 columns):

#	Column
0	email
1	full_name
2	how_was_your_last_month_in_your_assignment
3	How_Was_Your_Last_Month_Within_Your_Department
4	How_Was_Your_Last_Month_With_Us
5	What_Felt_Best_During_This_Last_Month
6	What_Should_We_Do_To_Make_You_Feel_Better_Satisfaction
7	Any_Additional_Suggestion_To_Improve
8	any_additional_comments
9	I_Would_Like_To_Get_Called_By
10	I_Would_Like_To_Provide_A_Copy_Of_My_Answers_To

dtypes: float64(3), int64(3), object(5)

memory usage: 8.7+ KB

Non-Null Count	Dtype
100 non-null	object
100 non-null	object
100 non-null	int64
100 non-null	int64
100 non-null	int64
37 non-null	object
37 non-null	object
15 non-null	object
0 non-null	object
0 non-null	object
0 non-null	object

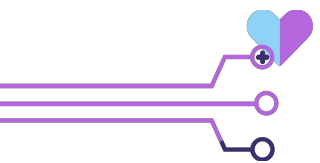
sdworxHRhackathon2021

Bestand Bewerken Weergeven Invoegen Opmaak Gegevens Extra Formulier Add-on's

100% € % .0 .00 123 Standaard ... 10 B I S

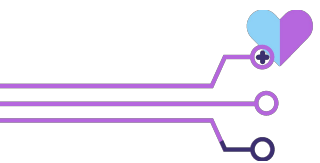
G6 less micro management

	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 impersonates me.
11	I would like that my assignments will be aligned with my career goals



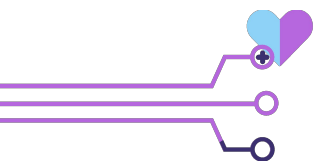
Entity extraction

	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	value all staff the same way	value	Skill
5	more transparency about the financial situatio...	financial	Skill



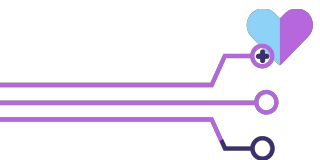
Sentiment extraction

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

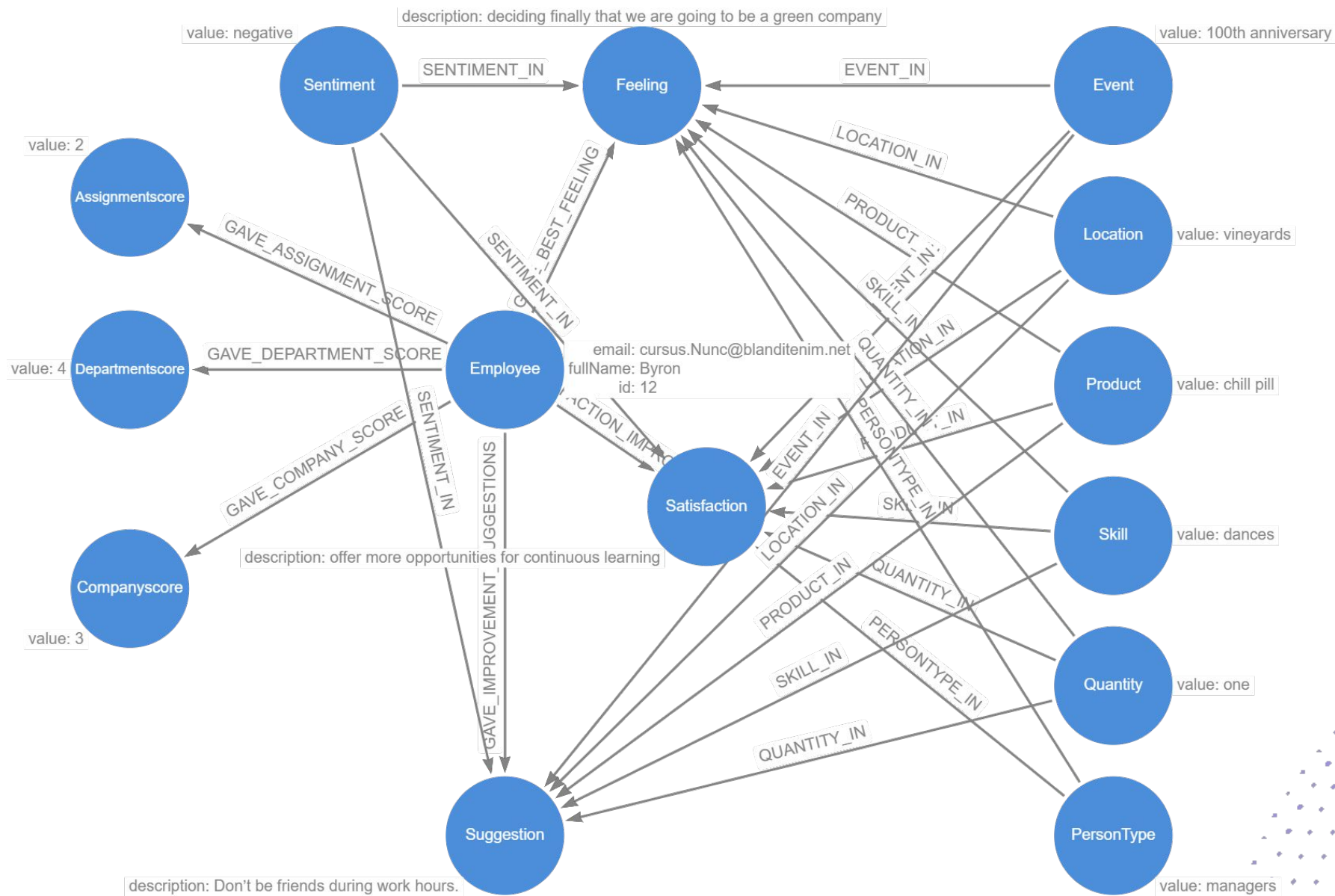


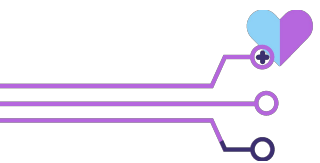
Key phrases extraction

	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 employees, making th...	['employees', 'partners', 'management decision'...
3	I don't feel that my skills and capacities are...	['skills', 'capacities', 'accounts']
4	less micro management	['micro management']

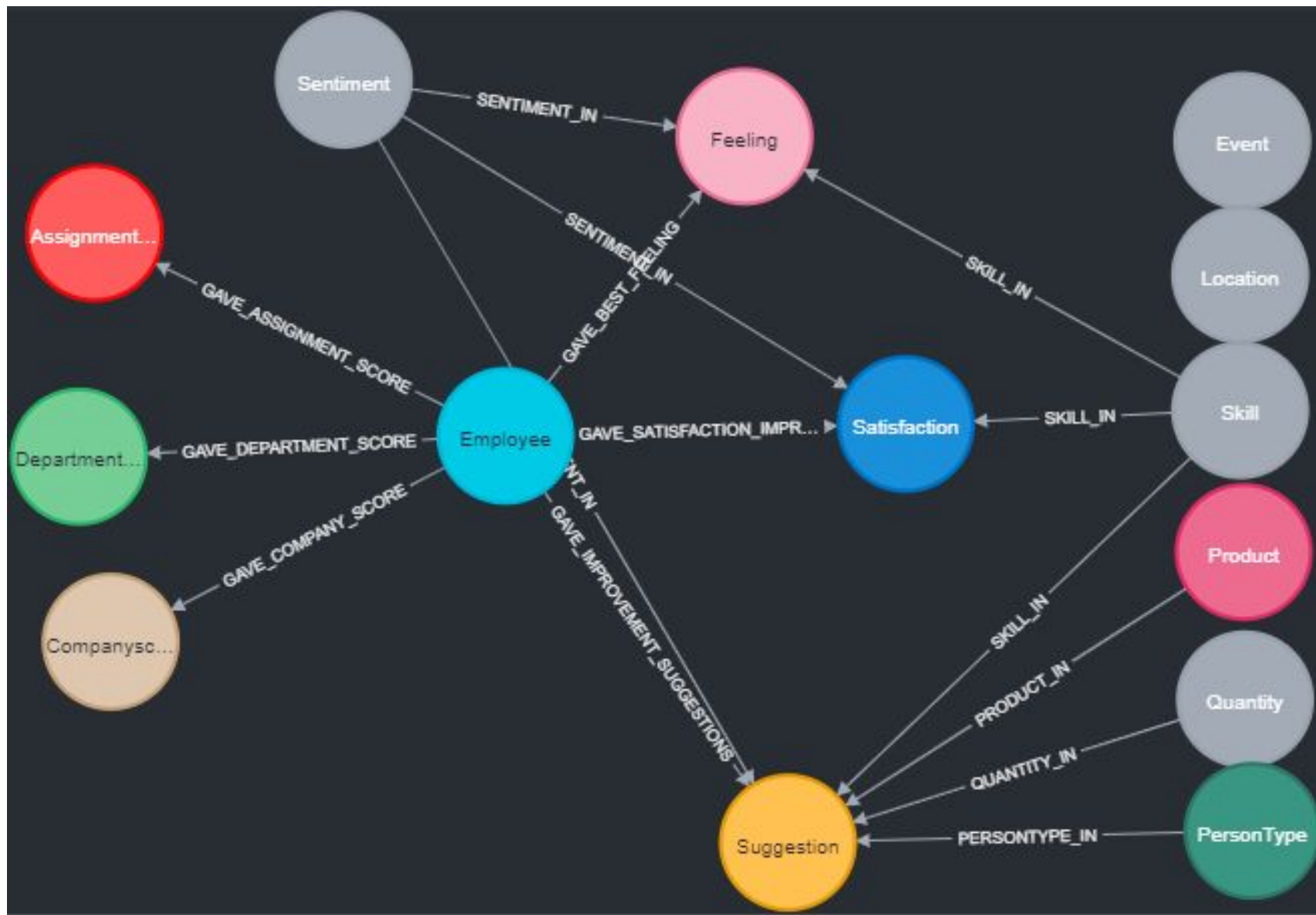


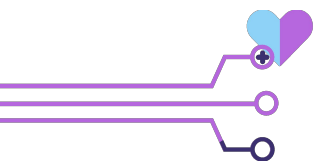
Datamodel





Graph database

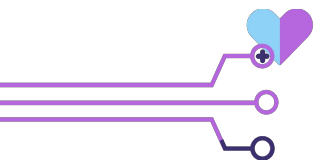




Some Cool Stuff We Did



what percentages has a positive, negative and neutral sentiment ?

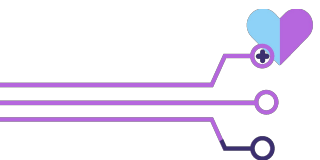


total sentiment percentage ?

```
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

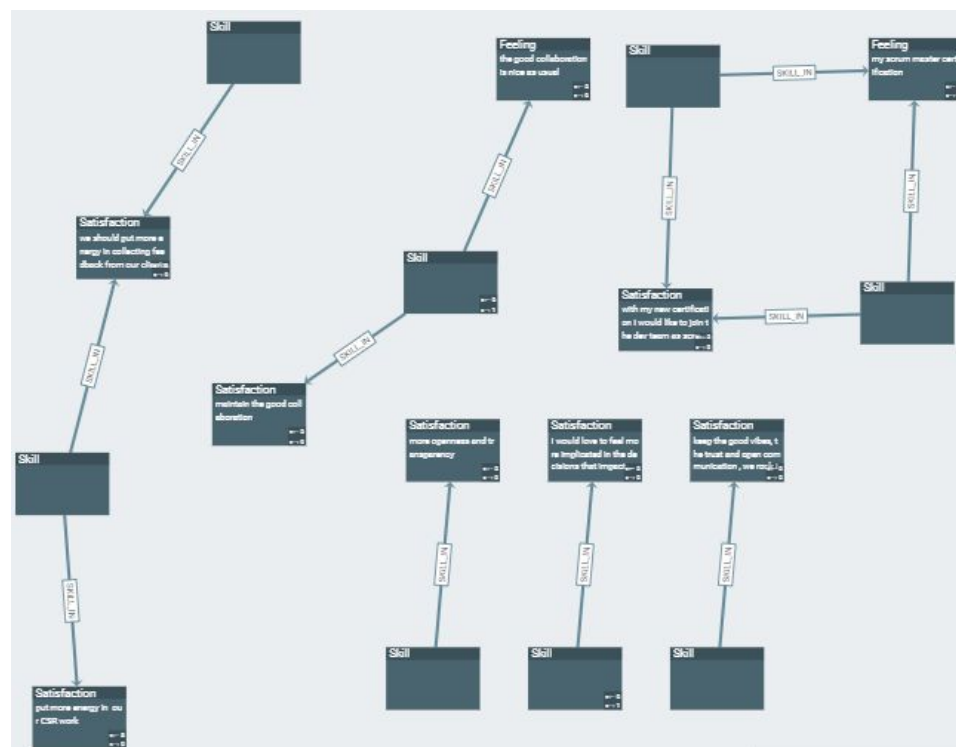
what percentages has a positive, negative and neutral sentiment ?



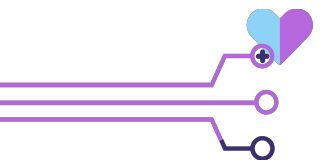
positive sentiment entity words ?

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

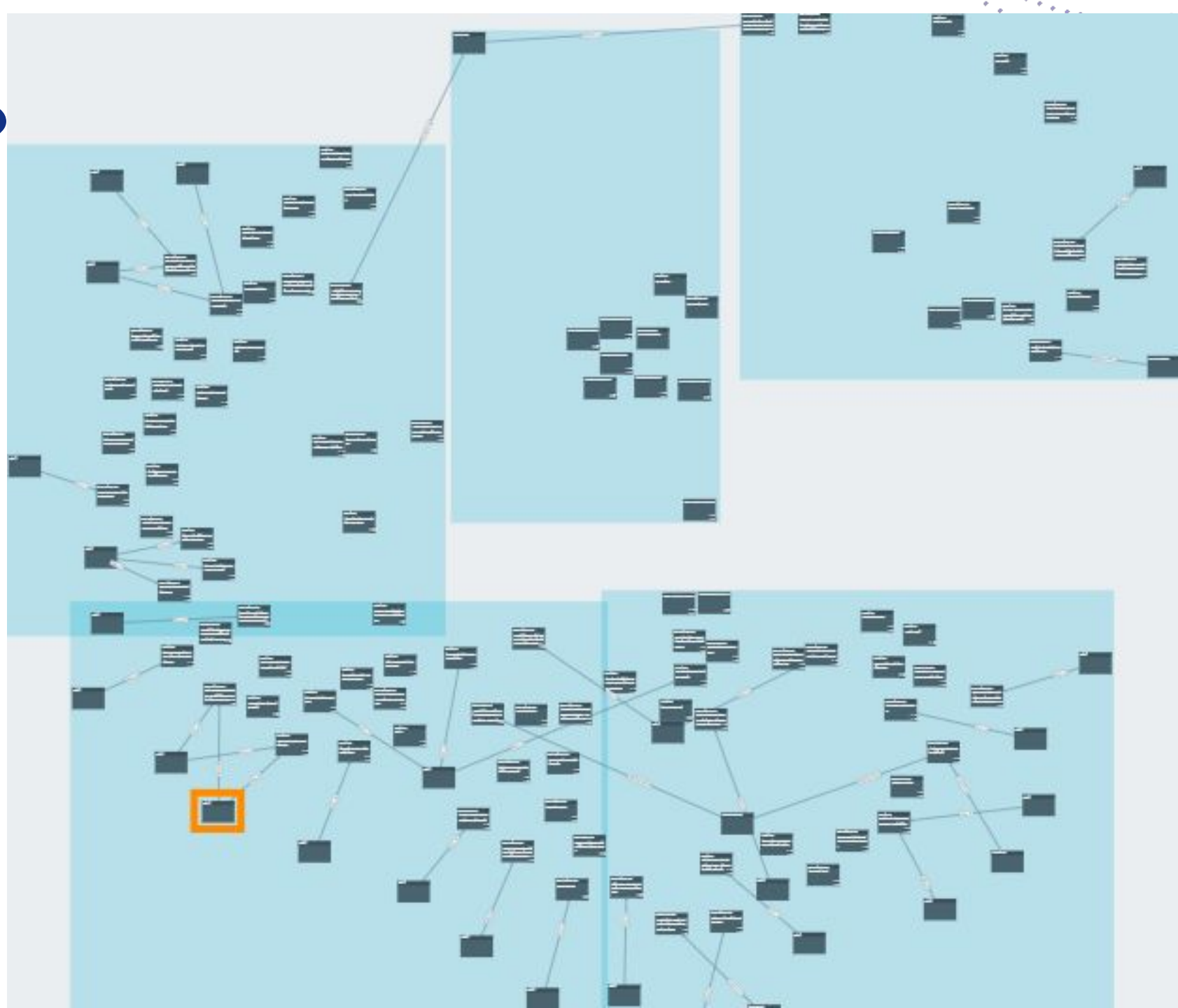
```
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->(feeling:Feeling)
WITH feeling
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->(satisfaction:Satisfaction)
WITH feeling, satisfaction
MATCH (s:Sentiment {value:'positive'})-[:SENTIMENT_IN]->(suggestion:Suggestion)
WITH feeling, satisfaction,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

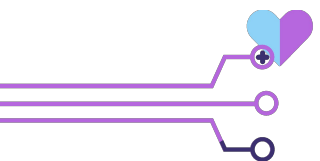


what percentages has a positive, negative and neutral sentiment ?



cluster ?



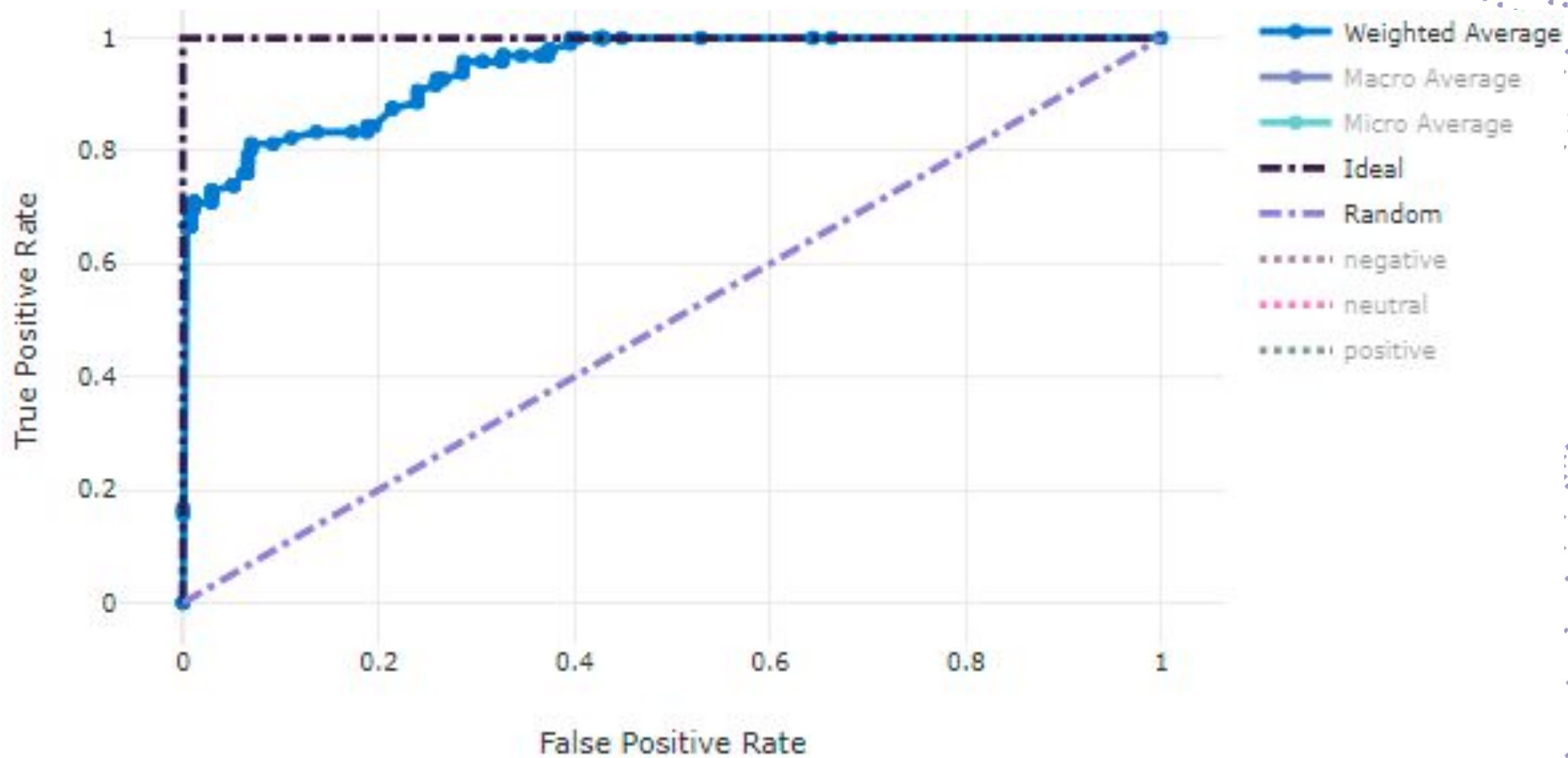


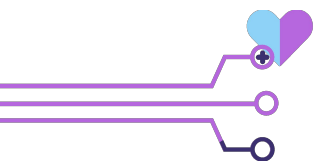
Machine Learning

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

accuracy
0.81

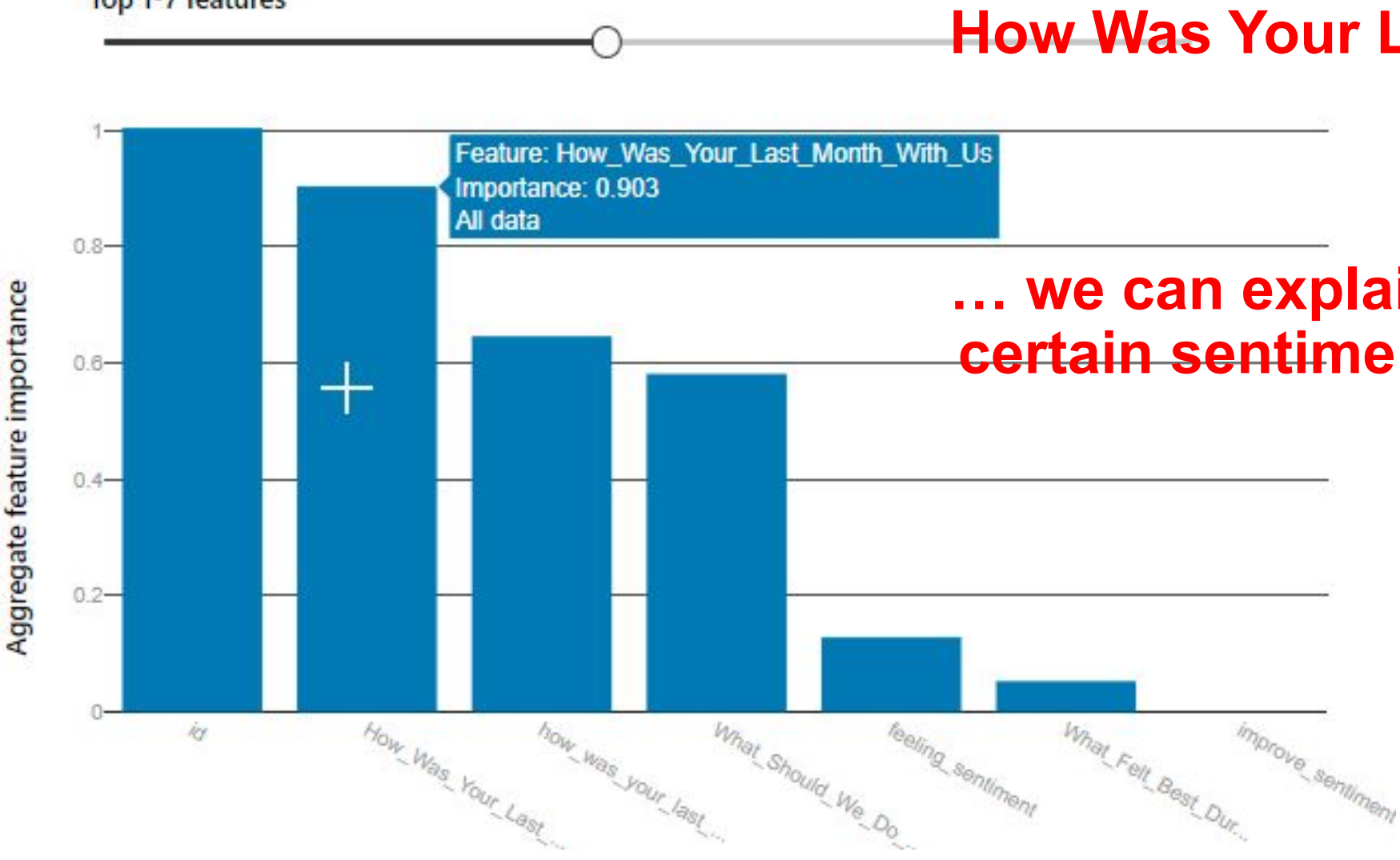
AUC_weighted
0.971





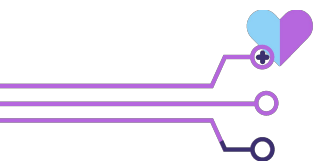
Most important Sentiment Influencer

Top 1-7 features

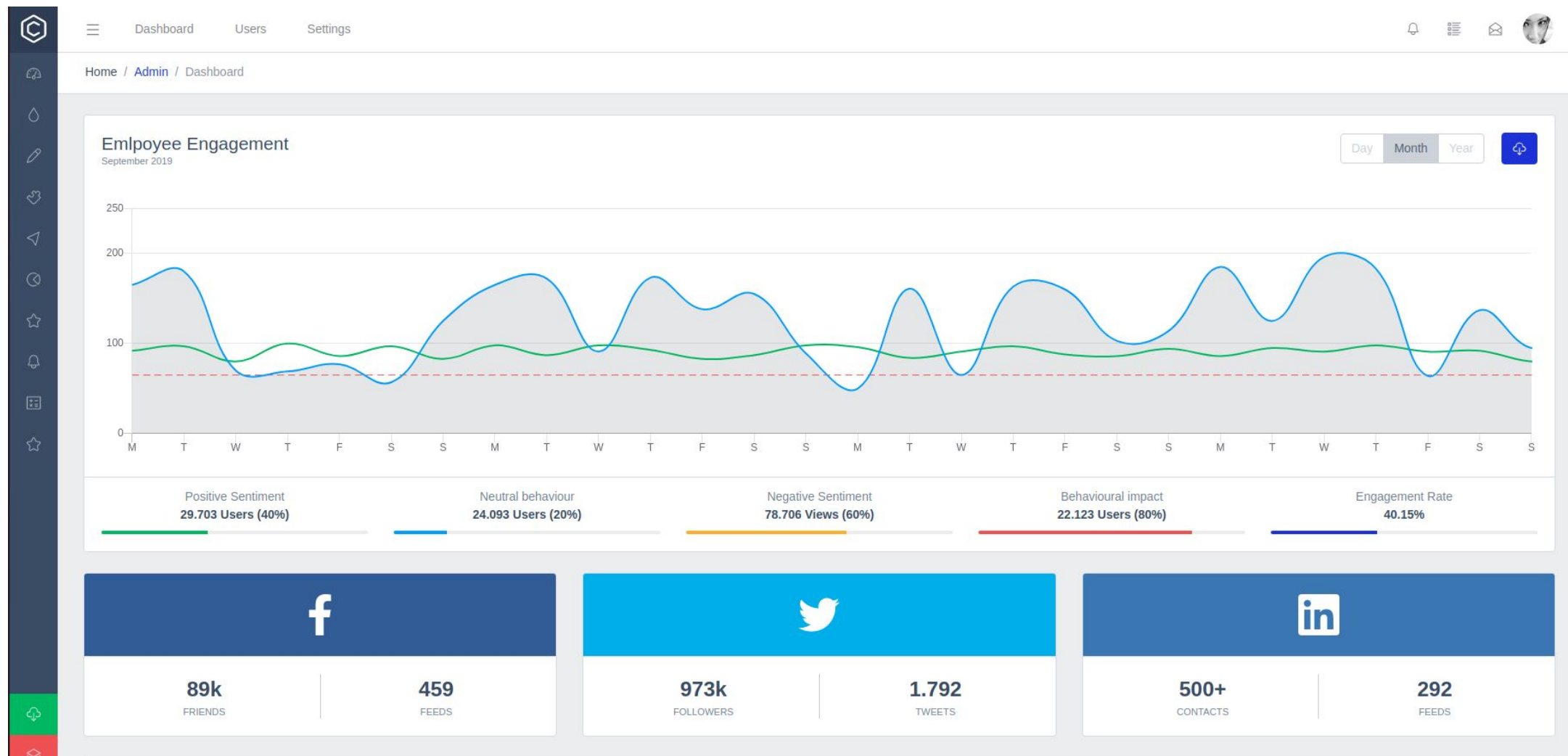


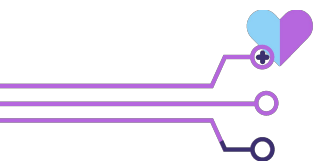
How Was Your Last Month With Us ?

... 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
```

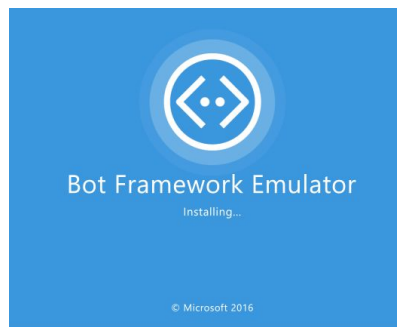
Bot framework Emulator

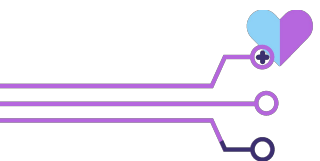
Entity Framework Core

Sql Server

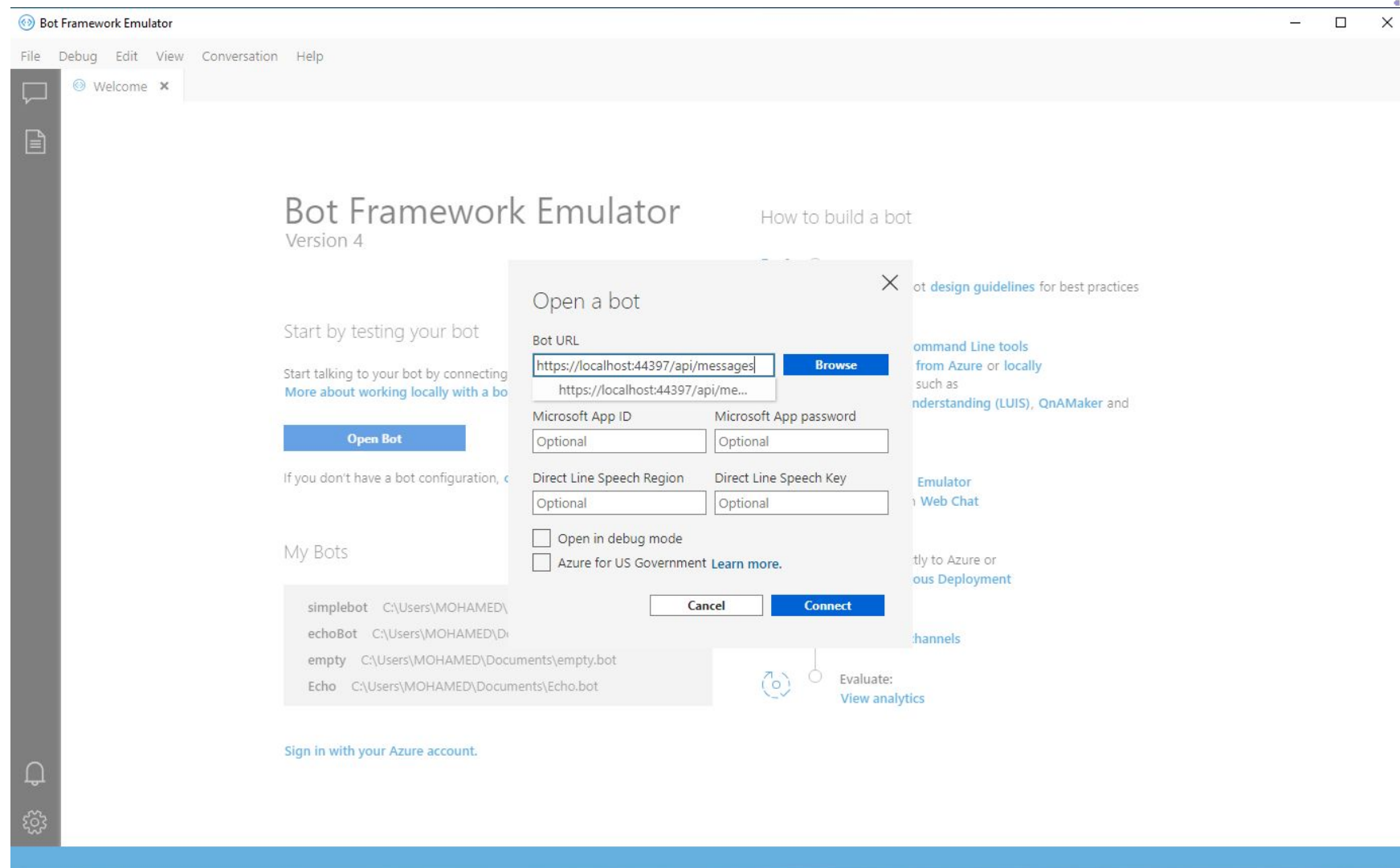
adaptivecards.io web site

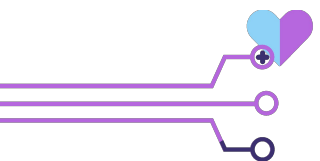
source code : [bayarmohamed/Hackathon \(github.com\)](https://github.com/bayarmohamed/Hackathon)





Bot Framework Emulator





ChatBot For Survey

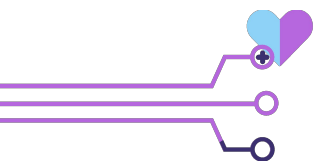
Use chatbot with
“Adaptive Cards”
SDK or json to
generate a survey
form

The screenshot displays the Bot Framework Emulator interface. The main chat area shows a survey form with the following fields:

- Your Email :
- How was your last month in your assignment
- How was your last month in your department
- How was your last month with us
- What felt best during this last month :
- What should we do to make you feel better and increase ...
- Any additional suggestion to improve
- Submit

The right-hand log panel shows the following activity log:

```
[19:31:30] Connecting to bot on https://localhost:44397/api/messages
[19:31:30] Emulator listening on http://[::]:51261
[19:31:30] ngrok not configured (only needed when connecting to remotely hosted bots)
[19:31:30] Connecting to bots hosted remotely
[19:31:30] Edit ngrok settings
[19:31:30] -> conversationUpdate
[19:31:31] POST 200 directline/conversations/<conversationId>/activities
[19:31:37] -> message hi
[19:31:37] <- message application/vnd.microsoft.card.adaptive
[19:31:37] POST 200 conversations/<conversationId>/activities/<activityId>
[19:31:37] POST 200 directline/conversations/<conversationId>/activities
```



Training Data with ML.Net

Select a binary classification as a scenario

The screenshot shows the ML.NET Model Builder interface with the 'Select a scenario' step selected in the left sidebar. The main area displays four scenarios: Text classification, Value prediction, Image classification, and Recommendation. The 'Text classification' scenario is highlighted with a red box. Below these are 'Limited scenarios' including Anomaly detection, Forecasting, Clustering, and Object detection.

1. Scenario | 2. Environment | 3. Data | 4. Train | 5. Evaluate | 6. Code

Select a scenario

Train with your data

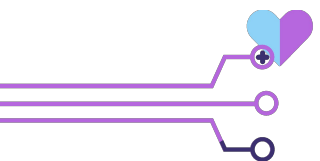
The following scenarios use Automated ML to train and pick the best model for your data. [Learn more about training with your own data in Model Builder.](#)

- Text classification**
Classify text data into 2+ categories, e.g. predict if comments are positive or negative sentiments.
[Local ML](#)
- Value prediction**
Predict a numeric value from your data (regression), e.g. predict the price of a house based on features like size, location, etc.
[Local ML](#)
- Image classification**
Classify images into 2+ categories, e.g. predict whether an image is of a dog or a cat.
[Azure ML](#) | [Local ML](#)
- Recommendation**
Produce a list of suggested items for a particular user, e.g. recommend products.
[Local ML](#)

Limited scenarios

The following scenarios are not yet supported by Automated ML, so walkthroughs with an example dataset and pre-defined training code are provided. [Learn more about examples in Model Builder.](#)

- Anomaly detection**
Detect abnormalities or outliers in data. This example detects spikes in shampoo sales.
[Example](#)
- Forecasting**
Predict future values based on previously observed time series values. This example predicts bike rental service demand.
[Example](#)
- Clustering**
Identify groups of related items without any pre-existing labels or categories. This example divides a set of iris flowers into different groups based on the flower's characteristics.
[Example](#)
- Object detection**
Detect and identify objects in images. This example detects objects (such as boats, people, and sofas) in images and draws bounding boxes around them.
[Example](#)



Training Data

data.csv - Excel

FICHIER ACCUEIL INSERTION MISE EN PAGE FORMULES DONNÉES RÉVISION AFFICHAGE TEAM

Calibri 11 A A Renvoyer à la ligne automatiquement Standard

Coller Presse-papiers Police Alignement Nombre Mise en conditionne

	A	B	C	D	E	F	G	H	I	J
1	email	full_name	how_was_y	How_Was_Y	How_Was_Y	What_Felt_E	What_Shoul	Any_Additional_Sugg	engagement	
2	Lorem.ipsun	Harrison	4	4	4	collaborating	keep the go	more sync between t	engaged	
3	elit.elit.ferr	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@niisini	Kane	4	4	3	getting to w	less micro m	I would like to talk to	engaged	
7	nisi@libero.	Kasper	4	4	4	being promc	more consid	Everything is all good	engaged	
8	dictum.eu@	Chase	2	2	2	not much for	keep the go	Fire half the staff and	notengaged	
9										
10										

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

ML.NET Model Builder - SDWorksChatBot.csproj

1. Scenario
2. Environment
3. Data
4. Train
5. Evaluate
6. Code

Add data

In order to build a model, you must add data and choose your column to predict.
[How do I get sample datasets and learn more?](#)

Input

Choose input data source from either SQL Server or File:

File

Select a file: C:\Users\MOHAMED\Desktop\SD\ ...

Supported file formats: .csv, .tsv or .txt.

Column to predict (Label): engagement

Input Columns (Features): 6 of 8 columns selected

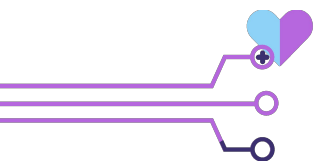
Data Preview

8 of 8 rows and 7 of 9 columns. (1 Label, 6 Features).

engagement (Label)	how_was_your_last_month_in_your_assignment	How_Was_Your_Last_Month_Within_Your_Department	How_Was_Your_Last_Month_With_Us	What_Felt_Best_During_This_Last_Month	Wh
engaged	4	4	4	collaborating closely with Damien, closing the deal with ucb	kee
notengaged	4	4	4	being promoted	ma
engaged	3	3	2	team building we had last month	mo
notengaged	1	1	1	not much for me	I dc
engaged	4	4	3	getting to work with you guys is a real pleasure	less
engaged	4	4	4	being promoted	mo
notengaged	2	2	2	not much for me	kee

Next step: train your model

Train



Train the model

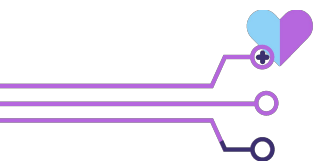
Train : generate a model to be used to predict survey
The training select the more adapted model (best model)
In our case the **AveragedPerceptronOva**.

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)

The screenshot displays the ML.NET Model Builder application window. On the left, a vertical sidebar lists six steps: 1. Scenario, 2. Environment, 3. Data, 4. Train (highlighted in blue), 5. Evaluate, and 6. Code. Each step is preceded by a green checkmark. The main area is titled 'Train' and contains the following information:

- A description: 'Specify a time to train for evaluating various models. [How long should I train for?](#)'
- A 'Training setup summary' dropdown menu.
- A 'Time to train (seconds):' input field with a value of '10' and an information icon.
- A 'Train again' button and a 'Training complete' status with a green checkmark.
- A 'Training results' section with the following data:

Best accuracy:	50%
Best model:	AveragedPerceptronOva
Training time:	2,96 seconds
Models explored (total):	1
- A 'Next step: evaluate your model' instruction and an 'Evaluate' button.



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.

Your Email :
bayarmohamed@gmail.com

How was your last month in your assignment
Good ▼

How was your last month in your department
Great ▼

How was your last month with us
Great ▼

What felt best during this last month :
collaborating closely with Damien, closing the deal with ucb

What should we do to make you feel better and increase ...
keep the good vibes, the trust and open communication , we rock !

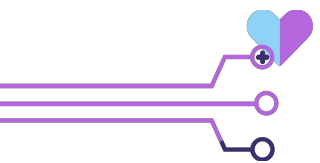
Any additional suggestion to improve
Everything is all good in the hood.

Submit

19 minutes ago

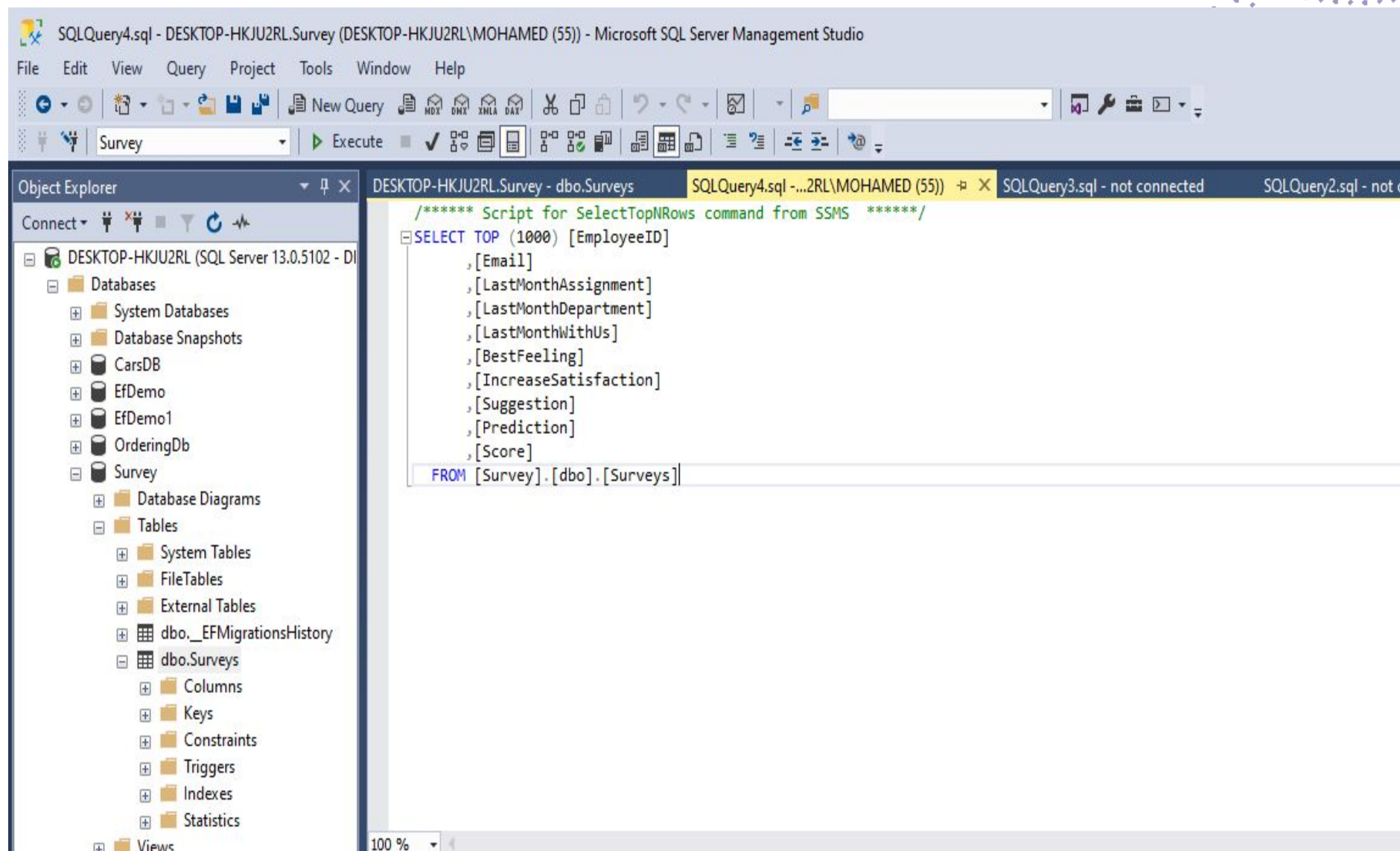
Engagment :engaged -> Score : 0.6473644

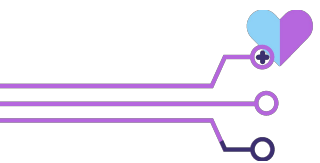
Just now



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

