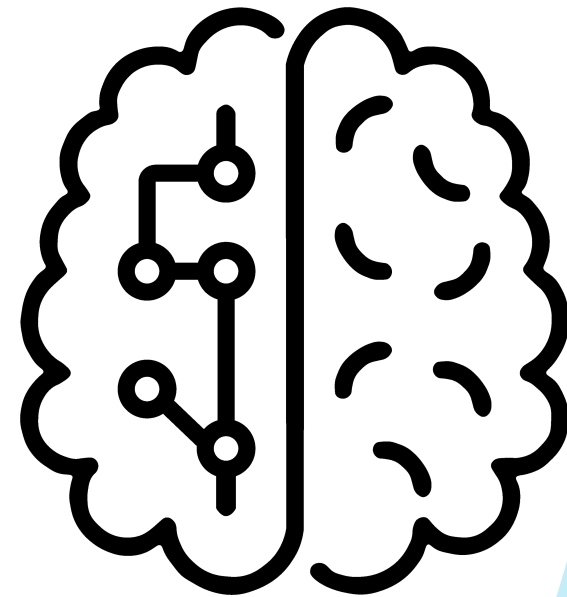
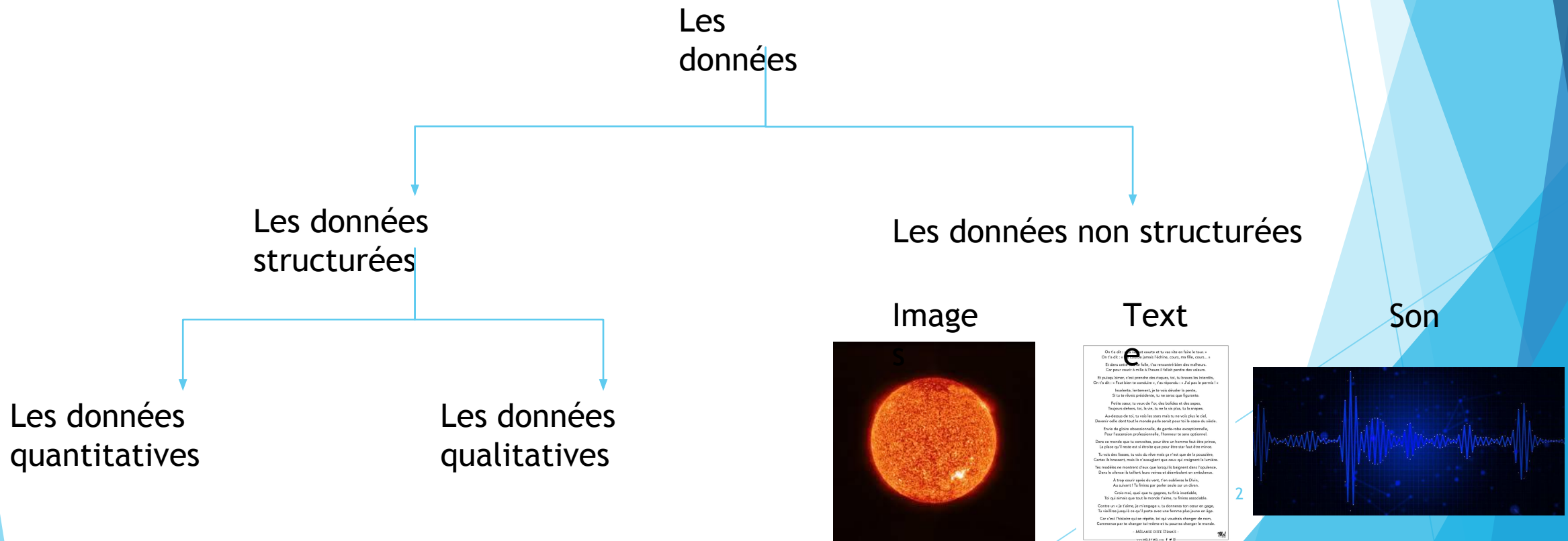


# Bag of words

AIForYou - Morgan Gautherot



# Les catégories de données



Text  
e

- MÉLANIE DITE DIAM'S -



## Données exploitables

## Modèles



# Le vecteur de vocabulaire

# D'une phrase à un sac de mots

Doc  
1

Le thé est  
bouillant



Traitements du  
texte

Doc  
1

[ 'thé', 'être',  
'bouillant' ]



Transformation en bag of  
words

# D'un corpus à un dataframe



Doc  
1

Tu me donnes envie de faire table rase du passé, et de t'aider à faire table rase du tien. Tu me donnes envie d'être pour toi ce que je n'ai jamais su être pour quelqu'un d'autre. Tu me donnes envie de t'offrir mon cœur, mes ambitions, mon amour sur un plateau d'argent. Tu me donnes envie de t'aimer à en faire trembler le monde entier. Tu me donnes envie de parler de sentiments et de bonheur tout en conjuguant mes verbes au présent. Tu me donnes envie de devenir quelqu'un qui t'apportera fierté et douceur. Tu me donnes envie de décrocher la Lune, de croire en l'impossible. Tu me donnes envie de vivre, bordel. Parce que tu es l'étincelle au fond de mes yeux, mon sourire terriblement sincère. Parce que je ne suis plus seule désormais, tu es là. Et c'est si précieux.

Doc  
2

Sophie fait les courses au supermarché avec son neveu Clément. Après avoir garé sa voiture sur le parking, elle a pris un chariot dans lequel Clément s'est installé. Tous deux glânent maintenant dans le magasin. Les clients sont nombreux à cause des promotions et des soldes. Sophie sort sa liste des courses pour ne rien oublier. Elle regarde les marques et compare les prix. Elle choisit des produits frais, des conserves et du poisson surgelé qu'elle met dans son chariot. Clément est turbulent, il s'agite sur son siège, il ne cesse de réclamer des bonbons, des biscuits, du chocolat, des jouets. Sophie le gronde gentiment puis elle regarde sa montre, il est déjà seize heures. Elle doit se hâter de rentrer car sa sœur elle est invitée avec Jean à dîner chez la maman de Clément, sa sœur Lucie. Sophie termine les courses, elle se dirige vers la caisse où elle attend calmement son tour. À la caisse, elle place ses achats sur le tapis roulant puis elle règle sa note avec sa carte bancaire. À présent, elle regagne le parking, cherche quelques instants sa voiture. Elle la repère au fond de l'allée puis se dirige vers son véhicule. Elle ouvre le coffre et y met les courses puis elle installe Clément dans son siège pour enfant. Elle part retrouver Jean et se préparer pour la soirée.

Le supermarché

Doc  
3

006  
L'AMITIÉ

L'amitié est un repère,  
une ligne d'horizon. Une lumière  
et une porte secrète vers les émotions...  
Elle nous aide par temps durs, elle fait rire,  
elle rassure. Elle ne juge pas, elle accepte,  
elle est là... Tantôt près, tantôt loin,  
mais toujours à distance de coeurs, notre  
amitié c'est du bonheur. Notre amitié, c'est  
une plante que je cultive avec grands soins,  
le trésor de mon petit jardin. Une pierre  
précieuse qui m'éclaire dans le noir, et  
quelqu'un qui connaît toute mon histoire...  
J'avance sur mon chemin, tu avances sur  
le tien. L'horizon devant nous est infini,  
tout comme notre amitié, merci.

Doc  
4

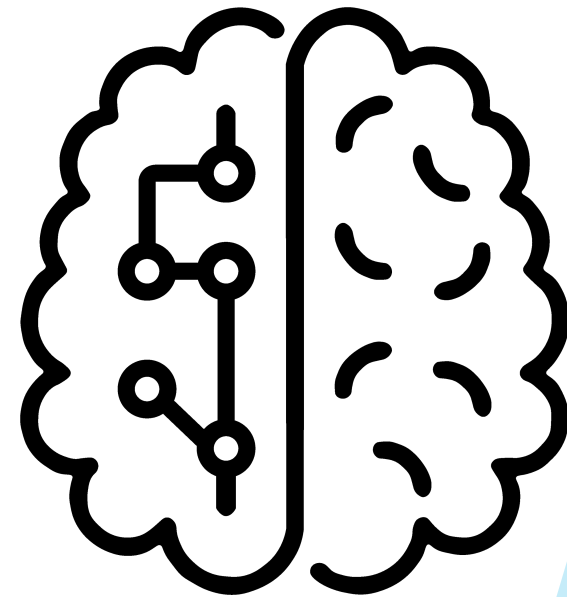
On l'a dit : « La vie est courte et tu vas vite en faire le tour. »  
On l'a dit : « Ne coude jamais l'histoire, aime, me lles, aime. »  
Et dans cette course folle, t'es rencontré bien des malheurs.  
Car pour courir à contre l'heure d'été j'ai perdu des valeurs.  
Et puis, hélas, c'est pendant des années, toi, la brève des années,  
On l'a dit : « Faut bien se conduire », t'es répondu : « J'ai pas le permis ! »  
Insolite, lentement, je te vais dévoiler la gente,  
Si tu te rends présente, tu ne sers que l'apogée.  
Petite sœur, tu veux de l'air, des bouillottes et des repas,  
Toujours effrayé, toi, le vie, tu ne te lais plus, tu te laisses.  
Au-dessus de toi, tu vois les stars mais tu ne vois plus le ciel,  
Devenir cette dexte tout le monde porte envie pour toi le casse du siècle.  
Entre de glorie obsolescence, de grande obsolescence,  
Pour l'accession professionnelle, l'honneur te sera optionnel.  
Dans ce monde que tu convoites, pour être un homme faut être prince,  
La place qu'il t'en aie et si droite que pour être une faut être reine.  
Tu vois des lasses, tu vois du rêve mais ça n'est que de la poussière,  
Certes tu brèves, mais tu n'imagines que ceux qui corrigent la lecture.  
Tu mérites me mérites d'être que lorsque tu brèves dans l'apogée.  
Dans le silence la laffent leurs vaines et déboulent en ambulance.  
À trop savoir après de vent, t'en mérites la Dexte,  
Au milieu ! Tu fesses pas partie mais tu en fais.  
Crois-moi, quoi que tu gagnes, tu fais inutilement.  
Toi qui crains que tout le monde t'aime, tu fesses inutilement.  
Contre un « je t'aime, je m'engage », tu donnes ton cœur en gage,  
Tu vas-t'en jusqu'à ce qu'il parte avec une femme plus jeune en âge.  
Car c'est l'histoire qui se réveille, toi qui voulais changer de route,  
Commence par te changer toi-même et tu pourras changer le monde.

- MÉLANIE DITE DINA -

V						
	0	1	0	0		0
	0	0	0	1		0
	0	1	0	0	0	0

# Le TF-IDF

AIForYou - Morgan Gautherot



# Définition du TF-IDF

TF = Term  
Frequency

IDF = Inverse Document  
Frequency



# Term Frequency

$$tf_{w,d} = \frac{n_{w,d}}{\sum_k n_{k,d}}$$

$d$  est un document de notre jeu de données

$w$  est un mot de notre document

$n_{w,d}$  est le nombre d'occurrences du mot  $w$  du document  $d$

# Inverse Document Frequency

$$idf_w = \log\left(\frac{|D|}{|\{d_i: t_w \in D\}|}\right)$$

$D$  désigne tous les documents de notre jeu de données

$w$  est un mot de notre document

$f(w, D)$  est le nombre de document  $D$  contenant le mot  $w$

# Calcul du TF-IDF

$$tfidf_{w,d} = tf_{w,d} * idf_w$$

$$tfidf_{w,d} = \frac{\text{fréquence du terme dans le document}}{\text{fréquence du terme dans le corpus}}$$

# Vecteur de vocabulaire

# D'une phrase à un vecteur

Doc  
1

Le thé est  
bouillant



Traitements du  
texte

Doc  
1

[ 'thé', 'être',  
'bouillant' ]



Transformation en  
vecteur

# D'un corpus à une matrice



Doc  
1

Tu me donnes envie de faire table rase du passé, et de t'aider à faire table rase du tien. Tu me donnes envie d'être pour toi ce que je n'ai jamais su être pour quelqu'un d'autre. Tu me donnes envie de t'offrir mon cœur, mes ambitions, mon amour sur un plateau d'argent. Tu me donnes envie de t'aimer à en faire trembler le monde entier. Tu me donnes envie de parler de sentiments et de bonheur tout en conjuguant mes verbes au présent. Tu me donnes envie de devenir quelqu'un qui t'apportera fierté et douceur. Tu me donnes envie de décrocher la Lune, de croire en l'impossible. Tu me donnes envie de vivre, bordel. Parce que tu es l'étincelle au fond de mes yeux, mon sourire terriblement sincère. Parce que je ne suis plus seule désormais, tu es là. Et c'est si précieux.

Doc  
2

Sophie fait les courses au supermarché avec son neveu Clément. Après avoir garé sa voiture sur le parking, elle a pris un chariot dans lequel Clément s'est installé. Tous deux glânent maintenant dans le magasin. Les clients sont nombreux à cause des promotions et des soldes. Sophie sort sa liste des courses pour ne rien oublier. Elle regarde les marques et compare les prix. Elle choisit des produits frais, des conserves et du poisson surgelé qu'elle met dans son chariot. Clément est turbulent, il s'agite sur son siège, il ne cesse de réclamer des bonbons, des biscuits, du chocolat, des jouets. Sophie le gronde gentiment puis elle regarde sa montre, il est déjà seize heures. Elle doit se hâter de rentrer car sa sœur elle est invitée avec Jean à dîner chez la maman de Clément, sa sœur Lucie. Sophie termine les courses, elle se dirige vers la caisse où elle attend calmement son tour. À la caisse, elle place ses achats sur le tapis roulant puis elle règle sa note avec sa carte bancaire. À présent, elle regagne le parking, cherche quelques instants sa voiture. Elle la repère au fond de l'allée puis se dirige vers son véhicule. Elle ouvre le coffre et y met les courses puis elle installe Clément dans son siège pour enfant. Elle part retrouver Jean et se préparer pour la soirée.

Le supermarché

Doc  
3

006  
L'AMITIÉ

L'amitié est un repère,  
une ligne d'horizon. Une lumière  
et une porte secrète vers les émotions...  
Elle nous aide par temps durs, elle fait rire,  
elle rassure. Elle ne juge pas, elle accepte,  
elle est là... Tantôt près, tantôt loin,  
mais toujours à distance de coeurs, notre  
amitié c'est du bonheur. Notre amitié, c'est  
une plante que je cultive avec grands soins,  
le trésor de mon petit jardin. Une pierre  
précieuse qui m'éclaire dans le noir, et  
quelqu'un qui connaît toute mon histoire...  
J'avance sur mon chemin, tu avances sur  
le tien. L'horizon devant nous est infini,  
tout comme notre amitié, merci.

Doc  
4

On t'a dit : « La vie est courte et tu vas vite en faire le tour. »  
On t'a dit : « Ne courais jamais l'histoire, cours, moi fille, cours. »  
Et dans cette course folle, t'as rencontré bien des malheurs.  
Car pour courir à cette vitesse t'as dû perdre bien des valeurs.  
Et puis, l'attente, c'est perdre des temps, toi, la bricoleuse des instants.  
On t'a dit : « Faut bien le conduire », t'as répondu : « J'ai pas le permis ! »  
Insolente, maintenant, je te vais diviser la gente.  
Si tu te dirais présidente, tu ne serais que l'égérie.  
Petite sœur, tu veux de l'or, des bouillottes et des sapins.  
Toujours effronté, toi, le vic, tu ne te lais plus le diol.  
Au-dessus de toi, tu vois les stars mais tu ne vois plus le ciel.  
Devenir ce que d'être tout le monde porte-moi pour toi le casse du siècle.  
Entre de glorie obsolescence, de grande obsolescence.  
Pour l'accession professionnelle, l'honneur te sera optionnel.  
Dans ce monde que tu convoites, pour être un homme faut être prince.  
La reine qu'il t'en a dit si droite que pour être une fille doit être reine.  
Tu vois des liesses, tu vois du rêve mais ça n'est que de la poussière.  
Certes tu brèves, mais tu n'imagines pas ceux qui corrigent la brève.  
Tu mérites un moment d'aise que l'orgueil du langage dans l'ambulance.  
Dans le silence la hantise leurs vaines et déboulent en ambulance.  
À trop savoir après de vent, t'en mérites le Divin.  
Au-dessus Tu feras pas partie mais ser un diabol.  
Crois-moi, quoi que tu gagnes, tu feras l'instabilité.  
Toi qui crâmes que tout le monde t'aime, tu feras l'obsolescence.  
Contre un « je t'aime, je m'engage », tu donneras ton cœur en gage.  
Tu verras jusqu'à ce qu'il parte avec une femme plus jeune en âge.  
Car c'est l'histoire qui se répète, toi qui vas tout changer de route.  
Commence par te changer toi-même et tu pourras changer le monde.

- MÉLANIE DITE DINA -

V						
	0		0	0		0
	0	0	0			0
	0					

# Notebook 2 - Implémentation du TF-IDF

Objectifs :

- ▶ Implémenter le bag of word pour un corpus de texte
- ▶ Implémenter le TF-IDF pour un corpus de texte