

Tenter la *Fake News Challenge*

Une proposition d'utilisation de la *Stance Detection* pour mieux anticiper le phénomène des *Fake News*

Enzo Poggio

Jury de mémoire:
Professeure Paola Merlo
Docteur Gabriele Musillo

Jeudi 28 Juin 2018

Sommaire

- 1 Fake News ?
 - Concept, Raisons, Origines
 - Propagations, Risques
 - Détections

Sommaire

1 Fake News ?

- Concept, Raisons, Origines
- Propagations, Risques
- Détections

2 Stance Detection

- Définition
- SemEval 2016 Task 6
- Fake News Challenge

Sommaire

- 1 Fake News ?
 - Concept, Raisons, Origines
 - Propagations, Risques
 - Détections
- 2 Stance Detection
 - Définition
 - SemEval 2016 Task 6
 - Fake News Challenge
- 3 Proposition FNC
 - Analyses et hypothèses
 - Combinaison : vote de majorité
 - Combinaison : 50/50 weighted average
 - Ensemble Learning : Single Layer Learner

Sommaire

- 1 Fake News ?
 - Concept, Raisons, Origines
 - Propagations, Risques
 - Détections
- 2 Stance Detection
 - Définition
 - SemEval 2016 Task 6
 - Fake News Challenge
- 3 Proposition FNC
 - Analyses et hypothèses
 - Combinaison : vote de majorité
 - Combinaison : 50/50 weighted average
 - Ensemble Learning : Single Layer Learner
- 4 Conclusion

Différents éléments de définitions

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Exemple de Fake News

« Alain Finkielkraut : « À ma mort, je léguerais tout mon patrimoine aux migrants. » » D'après *Secret News*.

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Exemple de Fake News

« Alain Finkielkraut : « À ma mort, je léguerais tout mon patrimoine aux migrants. » » D'après *Secret News*.

À ne pas confondre avec :

- Les erreurs journalistiques
- Les erreurs scientifiques
- Les parodies / satires

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Exemple de Fake News

« Alain Finkielkraut : « À ma mort, je léguerais tout mon patrimoine aux migrants. » » D'après *Secret News*.

Raisons principales

- La liberté d'expression sur le web
- Le **clickBait**
- Les idéologies

À ne pas confondre avec :

- Les erreurs journalistiques
- Les erreurs scientifiques
- Les parodies / satires

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Exemple de Fake News

« Alain Finkielkraut : « À ma mort, je léguerais tout mon patrimoine aux migrants. » » D'après *Secret News*.

À ne pas confondre avec :

- Les erreurs journalistiques
- Les erreurs scientifiques
- Les parodies / satires

Raisons principales

- La liberté d'expression sur le web
- Le **clickBait**
- Les idéologies

Exemple Fake News

- Vélès en Macédoine = Usine à Fake News
- Le Pizzagate

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Exemple de Fake News

« Alain Finkielkraut : « À ma mort, je léguerais tout mon patrimoine aux migrants. » » D'après *Secret News*.

À ne pas confondre avec :

- Les erreurs journalistiques
- Les erreurs scientifiques
- Les parodies / satires

Raisons principales

- La liberté d'expression sur le web
- Le **clickBait**
- Les idéologies

Exemple Fake News

- Vélès en Macédoine = Usine à Fake News
- Le Pizzagate

Origines médiatiques

La **négligence**, l'**unicité** et la **crise** des médias font que les Fake News peuvent se propager facilement.

Différents éléments de définitions

Définition primaire

Les **informations** qui ont pour but de **volontairement tromper** sont ce que l'on appelle des **Fake News**.

Exemple de Fake News

« Alain Finkielkraut : « À ma mort, je léguerais tout mon patrimoine aux migrants. » » D'après *Secret News*.

À ne pas confondre avec :

- Les erreurs journalistiques
- Les erreurs scientifiques
- Les parodies / satires

Raisons principales

- La liberté d'expression sur le web
- Le **clickBait**
- Les idéologies

Exemple Fake News

- Vélès en Macédoine = Usine à Fake News
- Le Pizzagate

Origines médiatiques

La **négligence**, l'**unicité** et la **crise** des médias font que les Fake News peuvent se propager facilement.

Organisations malveillantes

4chan et **reddit** sont en partie responsables de la **création de rumeurs**.

File: [tiphaine.macron.jpg](#) (424 KB, 830x508)

Anonymous (ID: [anonymous](#)) 12/04/2417 [Mon] 04:17:08 No.122527952 [Reply]



We need to stop this creep Macron from cuckooing up Europe, Le Pen la France's, and by extension, Europe's, last hope. We've got two weeks to wipe that Rothschild-smile off his face.

The left loves him because he yelled up his old teacher, Brigitte, who is like 30 years older than him, it plays into their 'truth speak, older woman can get hot young boyz' fantasy. Brigitte has got a 24 year old daughter from a previous marriage Tiphaine Auclerc. Wouldn't it be funny if I got out that Macron had engaged, or tried to engage in improper conduct with the young (arguably) hot daughter of his strong semi-serious wife in the past - that his wife is a woman scorned, but hush money from [the establishment] was doused on the matter to keep it a secret. (that last part about the hush money might be a stretch, but hey)

There have already been allegations about Macron being a closeted homosexual on the idea that he is a sexual pervert is already planted in the peoples' heads... people might buy it and that would land him into oblivion.

Well what do I know anon, I'm just a Swedenbuck grasping for straws here but something must be done in these coming two weeks and this is the best angle I came up with for now, the 'rogue-anon' would obviously have to help spread this in fog.

Rapides et dangereuses

Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Fake News sur les réseaux

- 38% en Français s'informent sur les réseaux
- 59% des liens partagés sur Twitter n'ont jamais été cliqués

Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Fake News sur les réseaux

- 38% en Français s'informent sur les réseaux
- 59% des liens partagés sur Twitter n'ont jamais été cliqués

Biais de confirmation et Intuition

A priori ou *a posteriori* les humains ont une faculté de juger assez mauvaises.

Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Fake News sur les réseaux

- 38% en Français s'informent sur les réseaux
- 59% des liens partagés sur Twitter n'ont jamais été cliqués

Biais de confirmation et Intuition

A priori ou *a posteriori* les humains ont une faculté de juger assez mauvaises.

Croyances et Préjugés

"Tous les cygnes sont blancs" ou racisme

Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Fake News sur les réseaux

- 38% en Français s'informent sur les réseaux
- 59% des liens partagés sur Twitter n'ont jamais été cliqués

Biais de confirmation et Intuition

A priori ou *a posteriori* les humains ont une faculté de juger assez mauvaises.

Croyances et Préjugés

"Tous les cygnes sont blancs" ou racisme

Risques politiques

Nier le consensus scientifique pour faire des choix politiques douteux.



Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Fake News sur les réseaux

- 38% en Français s'informent sur les réseaux
- 59% des liens partagés sur Twitter n'ont jamais été cliqués

Biais de confirmation et Intuition

A priori ou *a posteriori* les humains ont une faculté de juger assez mauvaises.

Croyances et Préjugés

"Tous les cygnes sont blancs" ou racisme

Risques politiques

Nier le consensus scientifique pour faire des choix politiques douteux.



Risques sanitaires

Les Fake News sont souvent liées sur les pseudo-effets secondaires de certains traitements.

Rapides et dangereuses

Propagations par les réseaux sociaux

Les réseaux sociaux sont la plus grosse plateforme d'échange de Fake News.

Fake News sur les réseaux

- 38% en Français s'informent sur les réseaux
- 59% des liens partagés sur Twitter n'ont jamais été cliqués

Biais de confirmation et Intuition

A priori ou *a posteriori* les humains ont une faculté de juger assez mauvaises.

Croyances et Préjugés

"Tous les cygnes sont blancs" ou racisme

Risques politiques

Nier le consensus scientifique pour faire des choix politiques douteux.



Risques sanitaires

Les Fake News sont souvent liées sur les pseudo-effets secondaires de certains traitements.

Vaccin Rougeole = Autisme

Le manque de couverture vaccinale peut faire revenir des maladies éradiquées et mortelles.

Humaine vs. Informatique

Humaine vs. Informatique

Comment vérifier des informations ?

Humaine vs. Informatique

Comment vérifier des informations ?

Avant de partager,

Il faut se questionner sur ce qui est raconté dans l'article et vérifier les sources. On est responsable de ce que l'on partage.

Humaine vs. Informatique

Comment vérifier des informations ?

Avant de partager,

Il faut se questionner sur ce qui est raconté dans l'article et vérifier les sources. On est responsable de ce que l'on partage.

Est-ce une information ?

Il faut se poser différentes questions : est-ce que cela a un intérêt public ? Est-ce factuel ? Est-ce vérifié ? Cela permet de distinguer les avis et les rumeurs des informations.

Humaine vs. Informatique

Comment vérifier des informations ?

Avant de partager,

Il faut se questionner sur ce qui est raconté dans l'article et vérifier les sources. On est responsable de ce que l'on partage.

Est-ce une information ?

Il faut se poser différentes questions : est-ce que cela a un intérêt public ? Est-ce factuel ? Est-ce vérifié ? Cela permet de distinguer les avis et les rumeurs des informations.

Ce site est-il fiable ?

A-t-il un onglet « À propos » sur sa page ? Est-il parodique ? Quelles sont les sources de ce site ?

Humaine vs. Informatique

Comment vérifier des informations ?

Avant de partager,

Il faut se questionner sur ce qui est raconté dans l'article et vérifier les sources. On est responsable de ce que l'on partage.

Est-ce une information ?

Il faut se poser différentes questions : est-ce que cela a un intérêt public ? Est-ce factuel ? Est-ce vérifié ? Cela permet de distinguer les avis et les rumeurs des informations.

Ce site est-il fiable ?

A-t-il un onglet « À propos » sur sa page ? Est-il parodique ? Quelles sont les sources de ce site ?

Comment vérifier des informations automatiquement ?

Humaine vs. Informatique

Comment vérifier des informations ?

Avant de partager,

Il faut se questionner sur ce qui est raconté dans l'article et vérifier les sources. On est responsable de ce que l'on partage.

Est-ce une information ?

Il faut se poser différentes questions : est-ce que cela a un intérêt public ? Est-ce factuel ? Est-ce vérifié ? Cela permet de distinguer les avis et les rumeurs des informations.

Ce site est-il fiable ?

A-t-il un onglet « À propos » sur sa page ? Est-il parodique ? Quelles sont les sources de ce site ?

Comment vérifier des informations automatiquement ? Stance Detection

Stance Detection, c'est quoi ? Un exemple !

Stance Detection, c'est quoi ? Un exemple !

Définition

La Stance Detection est la méthode qui permet de déterminer la relation **R** entre un énoncé **E** et une cible **C** donnée (**R1 : est pour** ou **R2 : est contre** ou **R3 : discute** ou **R4 : est non-lié**).

Stance Detection, c'est quoi ? Un exemple !

Définition

La Stance Detection est la méthode qui permet de déterminer la relation **R** entre un énoncé **E** et une cible **C** donnée (**R1 : est pour** ou **R2 : est contre** ou **R3 : discute** ou **R4 : est non-lié**).

Cible **C** (Ensemble de phrase)

- C1 Adam vend des pommes.
- C2 Charles rencontre Adam.
- C3 Charles achète une pomme à Adam.

Stance Detection, c'est quoi ? Un exemple !

Définition

La Stance Detection est la méthode qui permet de déterminer la relation **R** entre un énoncé **E** et une cible **C** donnée (**R1 : est pour** ou **R2 : est contre** ou **R3 : discute** ou **R4 : est non-lié**).

Cible **C** (Ensemble de phrase)

- C1 Adam vend des pommes.
- C2 Charles rencontre Adam.
- C3 Charles achète une pomme à Adam.

Relation **R** avec les énoncés **E** suivants

- R1 Charles a acheté une pomme.
- R2 Charles n'a pas acheté une pomme.
- R3 Charles et Adam aiment les pommes.
- R4 Jean est à la plage.

Stance Detection, c'est quoi ? Un exemple !

Définition

La Stance Detection est la méthode qui permet de déterminer la relation **R** entre un énoncé **E** et une cible **C** donnée (**R1 : est pour** ou **R2 : est contre** ou **R3 : discute** ou **R4 : est non-lié**).

Cible **C** (Ensemble de phrase)

- C1 Adam vend des pommes.
- C2 Charles rencontre Adam.
- C3 Charles achète une pomme à Adam.

Énoncés sémantiquement complexes

- E1 Charles a acheté un fruit.
- E2 Charles a une grosse pomme d'Adam.

Relation **R** avec les énoncés **E** suivants

- R1 Charles a acheté une pomme.
- R2 Charles n'a pas acheté une pomme.
- R3 Charles et Adam aiment les pommes.
- R4 Jean est à la plage.

Stance Detection, c'est quoi ? Un exemple !

Définition

La Stance Detection est la méthode qui permet de déterminer la relation **R** entre un énoncé **E** et une cible **C** donnée (**R1 : est pour** ou **R2 : est contre** ou **R3 : discute** ou **R4 : est non-lié**).

Cible **C** (Ensemble de phrase)

- C1 Adam vend des pommes.
- C2 Charles rencontre Adam.
- C3 Charles achète une pomme à Adam.

Énoncés sémantiquement complexes

- E1 Charles a acheté un fruit.
- E2 Charles a une grosse pomme d'Adam.

Relation **R** avec les énoncés **E** suivants

- R1 Charles a acheté une pomme.
- R2 Charles n'a pas acheté une pomme.
- R3 Charles et Adam aiment les pommes.
- R4 Jean est à la plage.

Genèse

L'article « Cats Rule and Dogs Drool ![...] » par Pranav Anand et al, (2011) tente de classer la prise de position dans des débats en ligne sur des sujets variés.

Description générale et sous-tâches

Description générale et sous-tâches

« Detecting Stance in Tweets »

Saif M. Mohammad et al (2016) nous propose trois classes possibles pour déterminer le parti pris (entre **favor**, **against** et **neither**) du tweeteur par rapport à son tweet grâce à la **Stance Detection**.

Description générale et sous-tâches

« Detecting Stance in Tweets »

Saif M. Mohammad et al (2016) nous propose trois classes possibles pour déterminer le parti pris (entre **favor**, **against** et **neither**) du tweeteur par rapport à son tweet grâce à la **Stance Detection**.

Exemple de tweet E pour la cible C

C Legalization of abortion

E A foetus has rights too ! Make your voice heard.

R **against**

Description générale et sous-tâches

« Detecting Stance in Tweets »

Saif M. Mohammad et al (2016) nous propose trois classes possibles pour déterminer le parti pris (entre **favor**, **against** et **neither**) du tweeteur par rapport à son tweet grâce à la **Stance Detection**.

Exemple de tweet E pour la cible C

C Legalization of abortion

E A foetus has rights too ! Make your voice heard.

R **against**

Évaluation

Évaluation sur le F1-score moyen est comparé aux baselines proposées par Saif M. Mohammad.

$$F_{avg} = \frac{F_{favor} + F_{against}}{2} \quad (1)$$

Description générale et sous-tâches

« Detecting Stance in Tweets »

Saif M. Mohammad et al (2016) nous propose trois classes possibles pour déterminer le parti pris (entre **favor**, **against** et **neither**) du tweeteur par rapport à son tweet grâce à la **Stance Detection**.

Exemple de tweet E pour la cible C

C Legalization of abortion

E A fetus has rights too ! Make your voice heard.

R **against**

Task A

La sous-tâche A est supervisée. Elle porte sur cinq différents sujets. Elle contient 4163 couples C/E labélisés R. On réserve 30% pour l'ensemble de test et le reste pour l'ensemble d'entraînement.

Évaluation

Évaluation sur le F1-score moyen est comparé aux baselines proposées par Saif M. Mohammad.

$$F_{avg} = \frac{F_{favor} + F_{against}}{2} \quad (1)$$

Description générale et sous-tâches

« Detecting Stance in Tweets »

Saif M. Mohammad et al (2016) nous propose trois classes possibles pour déterminer le parti pris (entre **favor**, **against** et **neither**) du tweeteur par rapport à son tweet grâce à la **Stance Detection**.

Exemple de tweet E pour la cible C

C Legalization of abortion

E A foetus has rights too ! Make your voice heard.

R **against**

Évaluation

Évaluation sur le F1-score moyen est comparé aux baselines proposées par Saif M. Mohammad.

$$F_{avg} = \frac{F_{favor} + F_{against}}{2} \quad (1)$$

Task A

La sous-tâche A est supervisée. Elle porte sur cinq différents sujets. Elle contient 4163 couples C/E labélisés R. On réserve 30% pour l'ensemble de test et le reste pour l'ensemble d'entraînement.

Task B

La sous-tâche B est non supervisée. Elle porte sur un seul thème ('Donald Trump'). L'ensemble test est constitué de 707 tweets E. Aucun ensemble d'entraînement n'a été donné. Mais un ensemble de 78 000 tweets non-labellisés à propos de Donald Trump était disponible.

Implémentations et Résultats

Implémentations et Résultats

Meilleur modèle Task A :
Réseau neuronal récurrent de MITRE

- 128 LSTM
- Vecteurs d'embedding de 256 dimensions (word2vec)

Implémentations et Résultats

Meilleur modèle Task A : Réseau neuronal récurrent de MITRE

- 128 LSTM
- Vecteurs d'embedding de 256 dimensions (word2vec)

Baseline	F_{favor}	$F_{against}$	F_{avg}
Majority class	52.01	78.44	65.22
SVM-unigrams	54.49	72.13	63.31
SVM-ngrams	62.98	74.98	68.98
SVM-ngrams-comb	54.11	70.01	62.06
Équipe			
MITRE	59.32	76.33	67.82
pkudblab	61.98	72.67	67.33
TakeLab	60.93	72.67	66.83

TABLE – Résultats pour la sous-tâche A des baselines et des 3 premiers participants ordonnés

Implémentations et Résultats

Meilleur modèle Task A : Réseau neuronal récurrent de MITRE

- 128 LSTM
- Vecteurs d'embedding de 256 dimensions (word2vec)

Meilleur modèle Task B : Réseau de neurones convolutionnels de pkudblab

- Matrice d'embedding (word2vec)
- Couche de convolution

Baseline	F_{favor}	$F_{against}$	F_{avg}
Majority class	52.01	78.44	65.22
SVM-unigrams	54.49	72.13	63.31
SVM-ngrams	62.98	74.98	68.98
SVM-ngrams-comb	54.11	70.01	62.06
Équipe			
MITRE	59.32	76.33	67.82
pkudblab	61.98	72.67	67.33
TakeLab	60.93	72.67	66.83

TABLE – Résultats pour la sous-tâche A des baselines et des 3 premiers participants ordonnés

Implémentations et Résultats

Meilleur modèle Task A : Réseau neuronal récurrent de MITRE

- 128 LSTM
- Vecteurs d'embedding de 256 dimensions (word2vec)

Baseline	F_{favor}	$F_{against}$	F_{avg}
Majority class	52.01	78.44	65.22
SVM-unigrams	54.49	72.13	63.31
SVM-ngrams	62.98	74.98	68.98
SVM-ngrams-comb	54.11	70.01	62.06
Équipe			
MITRE	59.32	76.33	67.82
pkudblab	61.98	72.67	67.33
TakeLab	60.93	72.67	66.83

TABLE – Résultats pour la sous-tâche A des baselines et des 3 premiers participants ordonnés

Meilleur modèle Task B : Réseau de neurones convolutionnels de pkudblab

- Matrice d'embedding (word2vec)
- Couche de convolution

Baseline	F_{favor}	$F_{against}$	F_{avg}
Majority class	0.00	59.44	29.72
SVM-ngrams-comb	18.42	38.45	28.43
Équipe			
pkudblab	57.39	55.17	56.28
LitisMind	30.04	59.28	44.66
INF-UFRGS	32.56	52.09	42.32

TABLE – Résultats pour la sous-tâche B des baselines et 3 premiers participants ordonnés

Description de la tâche

Description de la tâche

But

L'objectif du FNC est d'explorer comment les technologies d'intelligence artificielle pourraient être utilisées pour lutter contre les Fake News.

Description de la tâche

But

L'objectif du FNC est d'explorer comment les technologies d'intelligence artificielle pourraient être utilisées pour lutter contre les Fake News.

Données et origines des données

Le FNC est une tâche partagée supervisée. Les organisateurs définissent les données en termes d'entrées (C et E) et de sorties (R). Une entrée est donc formée par le couple d'une affirmation E (titre d'un article) et un corps de texte C (corps d'un article). Une sortie est la relation R corps du texte par rapport à la revendication faite dans l'affirmation.

Description de la tâche

But

L'objectif du FNC est d'explorer comment les technologies d'intelligence artificielle pourraient être utilisées pour lutter contre les Fake News.

Données et origines des données

Le FNC est une tâche partagée supervisée. Les organisateurs définissent les données en termes d'entrées (C et E) et de sorties (R). Une entrée est donc formée par le couple d'une affirmation E (titre d'un article) et un corps de texte C (corps d'un article). Une sortie est la relation R corps du texte par rapport à la revendication faite dans l'affirmation.

Labels R

- agree** : Le corps du texte **C** est en accord avec l'affirmation **E**.
- disagree** : Le corps du texte **C** n'est pas d'accord avec l'affirmation **E**.
- discuss** : Le corps du texte discute **C** le même sujet que l'affirmation **E**, mais ne prend pas parti.
- unrelated** : Le corps du texte **C** traite d'un sujet différent de l'affirmation **E**.

Description de la tâche

But

L'objectif du FNC est d'explorer comment les technologies d'intelligence artificielle pourraient être utilisées pour lutter contre les Fake News.

Données et origines des données

Le FNC est une tâche partagée supervisée. Les organisateurs définissent les données en termes d'entrées (C et E) et de sorties (R). Une entrée est donc formée par le couple d'une affirmation E (titre d'un article) et un corps de texte C (corps d'un article). Une sortie est la relation R corps du texte par rapport à la revendication faite dans l'affirmation.

Labels R

- agree** : Le corps du texte **C** est en accord avec l'affirmation **E**.
- disagree** : Le corps du texte **C** n'est pas d'accord avec l'affirmation **E**.
- discuss** : Le corps du texte discute **C** le même sujet que l'affirmation **E**, mais ne prend pas parti.
- unrelated** : Le corps du texte **C** traite d'un sujet différent de l'affirmation **E**.

Exemples d'évaluation

Classe Gold	Classe attribuée	Score
unrelated	unrelated	+0.25
agree	unrelated	+0
agree	disagree	+0.25
agree	agree	+0.75

TABLE — Score en fonction de l'attribution de classe. Related en rouge

Solat in the news (Notre baseline)

Solat in the news (Notre baseline)

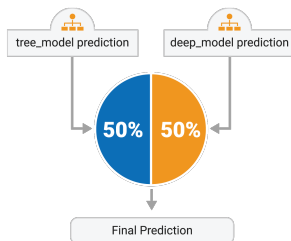
Mesure	Solat			
	agree	disagree	discuss	unrelated
Précision	0.59	0.02	0.76	0.99
Rappel	0.5	0.36	0.76	0.97
F1score	0.54	0.04	0.76	0.98
Exactitude	89.08			
Score FNC	9556.5			
Pourcentage FNC	82.02			

TABLE – Mesures pour le modèle : Solat

Solat in the news (Notre baseline)

Mesure	Solat			
	agree	disagree	discuss	unrelated
Précision	0.59	0.02	0.76	0.99
Rappel	0.5	0.36	0.76	0.97
F1score	0.54	0.04	0.76	0.98
Exactitude	89.08			
Score FNC	9556.5			
Pourcentage FNC	82.02			

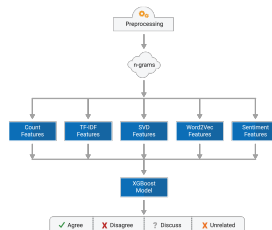
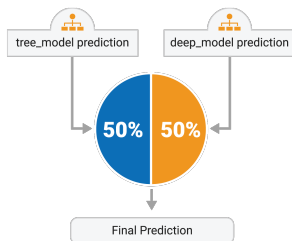
TABLE – Mesures pour le modèle : Solat



Solat in the news (Notre baseline)

Mesure	Solat			
	agree	disagree	discuss	unrelated
Précision	0.59	0.02	0.76	0.99
Rappel	0.5	0.36	0.76	0.97
F1score	0.54	0.04	0.76	0.98
Exactitude	89.08			
Score FNC	9556.5			
Pourcentage FNC	82.02			

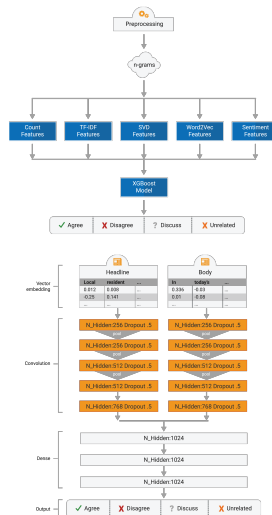
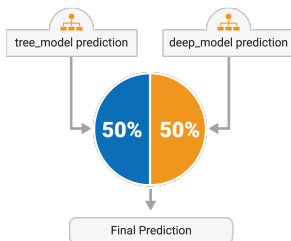
TABLE – Mesures pour le modèle : Solat



Solat in the news (Notre baseline)

	Solat			
Mesure	agree	disagree	discuss	unrelated
Précision	0.59	0.02	0.76	0.99
Rappel	0.5	0.36	0.76	0.97
F1score	0.54	0.04	0.76	0.98
Exactitude	89.08			
Score FNC	9556.5			
Pourcentage FNC	82.02			

TABLE – Mesures pour le modèle : Solat



Athene

Athene

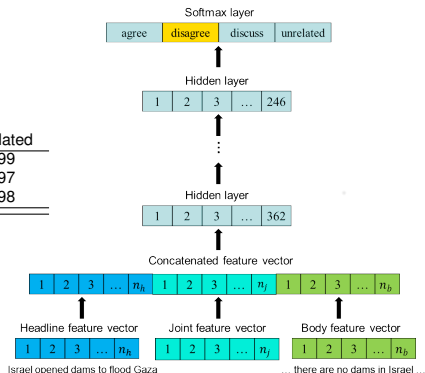
Mesure	Athene			
	agree	disagree	discuss	unrelated
Précision	0.45	0.09	0.81	0.99
Rappel	0.54	0.38	0.75	0.97
F1score	0.49	0.15	0.78	0.98
Exactitude	89.48			
Score FNC	9550.75			
Pourcentage FNC	81.97			

TABLE – Mesures pour le modèle : Athene

Athene

Mesure	Athene			
	agree	disagree	discuss	unrelated
Précision	0.45	0.09	0.81	0.99
Rappel	0.54	0.38	0.75	0.97
F1score	0.49	0.15	0.78	0.98
Exactitude	89.48			
Score FNC	9550.75			
Pourcentage FNC	81.97			

TABLE – Mesures pour le modèle : Athene



UCLmr

UCLmr

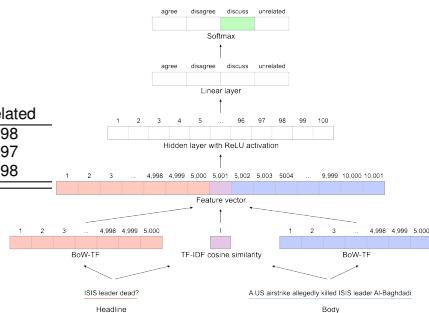
Mesure	UCL-mr			
	agree	disagree	discuss	unrelated
Précision	0.44	0.07	0.81	0.98
Rappel	0.53	0.43	0.69	0.97
F1score	0.48	0.11	0.75	0.98
Exactitude	88.46			
Score FNC	9521.5			
Pourcentage FNC	81.72			

TABLE – Mesures pour le modèle : UCL-mr

UCLmr

Mesure	UCL-mr			
	agree	disagree	discuss	unrelated
Précision	0.44	0.07	0.81	0.98
Rappel	0.53	0.43	0.69	0.97
F1score	0.48	0.11	0.75	0.98
Exactitude	88.46			
Score FNC	9521.5			
Pourcentage FNC	81.72			

TABLE – Mesures pour le modèle : UCL-mr



Voie à explorer

Voie à explorer

Analyses

- Confusion entre **agree** et **discuss**
- Les **disagree** sont sous-représentés, ils ne rapportent que peu de point.

Voie à explorer

Analyses

- Confusion entre **agree** et **discuss**
- Les **disagree** sont sous-représentés, ils ne rapportent que peu de point.

Hypothèses

- Mieux distinguer **agree** et **discuss** donne **des meilleurs résultats**
- Solat est premier en utilisant des une combinaison de modèles. L'**ensemble learning** ou des **modèles combinatoires** donnerait de **meilleurs résultats**.

The simplest, the better

The simplest, the better

Principe

Modèle Dominant	Modèle assujetti 1	Modèle assujetti 2	Classe prédite
agree	agree	unrelated	agree
disagree	discuss	discuss	discuss
discuss	disagree	agree	discuss

TABLE – Exemple du vote de majorité

The simplest, the better

Principe

Modèle Dominant	Modèle assujetti 1	Modèle assujetti 2	Classe prédite
agree	agree	unrelated	agree
disagree	discuss	discuss	discuss
discuss	disagree	agree	discuss

TABLE – Exemple du vote de majorité

Résultats

Mesure	agree	disagree	discuss	unrelated
Précision	0.48	0.07	0.84	0.99
Rappel	0.57	0.41	0.75	0.97
F1score	0.52	0.11	0.79	0.98
Exactitude	90.03			
Score FNC	9702.75			
Pourcentage FNC	83.28			

TABLE – Mesures pour le modèle : Vote de majorité Dominant Athene

The simplest, the better

Principe

Modèle Dominant	Modèle assujetti 1	Modèle assujetti 2	Classe prédite
agree	agree	unrelated	agree
disagree	discuss	discuss	discuss
discuss	disagree	agree	discuss

TABLE – Exemple du vote de majorité

Résultats

Mesure	agree	disagree	discuss	unrelated
Précision	0.48	0.07	0.84	0.99
Rappel	0.57	0.41	0.75	0.97
F1score	0.52	0.11	0.79	0.98
Exactitude	90.03			
Score FNC	9702.75			
Pourcentage FNC	83.28			

TABLE – Mesures pour le modèle : Vote de majorité Dominant Athene

Leçon d'humilité et compréhension du modèle

Ce modèle est la plus simple des combinaisons que l'on pouvait imaginer entre les modèles. Le vote de majorité est la seule de nos combinaisons à utiliser les résultats de 3 sous-modèles ce qui explique certainement son succès.

Plus de labels mais des probabilités...

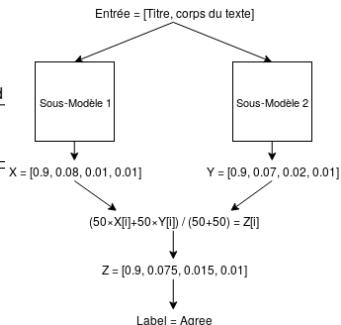
Plus de labels mais des probabilités...

Résultats

Meilleure combinaison probabiliste				
Mesure	agree	disagree	discuss	unrelated
Précision	0.53	0.0	0.81	0.99
Rappel	0.55	0.33	0.74	0.97
F1score	0.54	0.0	0.77	0.98
<hr/>				
Exactitude	89.65			
Score FNC	9633.25			
Pourcentage FNC	82.68			

TABLE – Mesures pour le modèle : mixte UCL-mr/Solat sans TF-Idf Moyenne

Principe



On rajoute une couche d'apprentissage !

On rajoute une couche d'apprentissage !

Résultats

Mixte UCL-mr/Solat sans TF-Idf SLL				
Mesure	agree	disagree	discuss	unrelated
Précision	0.56	0.02	0.79	0.98
Rappel	0.52	0.58	0.72	0.97
F1score	0.54	0.04	0.76	0.98
Exactitude	89.01			
Score FNC	9606.75			
Pourcentage FNC	82.45			

TABLE – Mesures pour le modèle : Mixte UCL-mr/Solat sans TF-Idf SLL

Modèle

- Vecteurs d'entrée : concaténation des probabilités des sous-modèles **mixtes UCL-mr/Solat sans TF-Idf Moyenne**.
- Couche neuronale de **32 unités**
- **Softmax** de sortie **4 unités** pour chaque label
- Corpus d'entraînement **croisé**

Discussion et conclusion

Discussion et conclusion

Discussion

Discussion et conclusion

Discussion

Hypothèses

Nous avons bien **vérifié nos hypothèses**.

Nous avons de meilleurs résultats :

- En **combinant** les résultats de **plusieurs systèmes gagnants de la FNC**
- En **distinguant** mieux les labels **agree** des **discuss**

Discussion et conclusion

Discussion

Hypothèses

Nous avons bien **vérifié nos hypothèses**.

Nous avons de meilleurs résultats :

- En **combinant** les résultats de **plusieurs systèmes gagnants de la FNC**
- En **distinguant** mieux les labels **agree** des **discuss**

Meilleur résultat

Le modèle de **vote de majorité où le système Athene est dominant** dépasse notre baseline de **1.26** avec le score de **83,28** au pourcentage FNC.

Conclusion

Discussion et conclusion

Discussion

Hypothèses

Nous avons bien **vérifié nos hypothèses**.

Nous avons de meilleurs résultats :

- En **combinant** les résultats de **plusieurs systèmes gagnants de la FNC**
- En **distingant** mieux les labels **agree** des **discuss**

Meilleur résultat

Le modèle de **vote de majorité** où le **système Athene est dominant** dépasse notre baseline de **1.26** avec le score de **83,28** au pourcentage FNC.

Conclusion

Objectifs du mémoire atteints

- Définir les Fake News **exhaustivement**.
- Présenter la **Stance Detection**.
- Dépasser **l'état de l'art** en détection de Fake News.

Discussion et conclusion

Discussion

Hypothèses

Nous avons bien **vérifié nos hypothèses**.
Nous avons de meilleurs résultats :

- En **combinant** les résultats de **plusieurs systèmes gagnants de la FNC**
- En **distingant** mieux les labels **agree** des **discuss**

Meilleur résultat

Le modèle de **vote de majorité** où le **système Athene est dominant** dépasse notre baseline de **1.26** avec le score de **83,28** au pourcentage FNC.

Conclusion

Objectifs du mémoire atteints

- Définir les Fake News **exhaustivement**.
- Présenter la **Stance Detection**.
- Dépasser **l'état de l'art** en détection de Fake News.

Détecteur de Fake News ?

Peut-on vraiment appeler notre modèle un **détecteur de Fake News** ? Nous n'avons ici qu'un **proto-détecteur de Fake News** qui n'est pas garant de la **véracité**.

Merci pour votre écoute