



# Classification d'assertions selon leur valeurs de véracité (automatic fact-checking)

Sabri BENBRAHIM - 21604014  
Yann DUFRESNE - 20055179

Bénédicte DAYNAC - 21605192  
Llivia LANGEVIN - 21604582

# Introduction



## Objectif:

Définir un modèle de classification textuelle pour vérifier la véracité d'une assertion

- Traitement des données
- Feature engineering
- Recherche de modèles de classifications

# Préparation des données



- Nettoyage des données "à la main" :
  - Corrections des erreurs de lignes et de colonnes
  - Normalisation des valeurs de rating à l'aide d'un script
  - Création de clé unique
  - Sélection des colonnes pertinentes
  - Division du jeu de données en deux (True vs False) et (True-False vs Mixture)

# Pré-traitements



- Nettoyage des claims
  - Encodage des caractères
  - Balises HTML
  - Contractions anglaises
- Nettoyage des tokens
  - Tout en minuscules
  - Nombres en lettres
  - Suppression des symboles particuliers
- Préparation au classifieur

# Impact des features sur la prédiction



- Suppression des stop words
- Lemmatisation
- Ajout des métadonnées
- Méthodes de resampling

# Impact des features sur la prédiction

	T vs F	T-F vs M	
Initial	48.497%	59.670%	Retenu
Suppression des Stop Words	46.514%	59.073%	
Lemmatisation	46.790%	58.753%	
Ajout des méta-données liées à l'auteur	<b>53.367%</b>	<b>64.100%</b>	
Ajout des autres méta-données	51.126%	63.712%	
Downsampling	64.748%	68.895%	
Upsampling	75.325%	73.025%	
Downresample	65.657%	70.064%	Retenu
Upresample	<b>75.318%</b>	<b>73.025%</b>	

Tableau basé sur les valeurs de F1-score

# Modèles de classification

Modèle de classification	True/False	True-False/Mixture
GaussianNB	77.488627%	72.252368%
MultinomialNB	77.395073%	73.395706%
LinearSVC	83.101483%	77.257629%
LogisticRegression	84.013906%	77.407122%
DecisionTreeClassifier	63.316127%	NA
RandomForestClassifier	63.787304%	NA
SGDClassifier	78.988118%	NA

# Bilan

---

- Utilisation de Collaboratory by Google
- **84%** via le classifieur LogisticRegression
- Améliorations possibles
  - Sélection de variable
  - N-Grams
  - Pos-Tagging

