

DeFT 2019

Recherche et extraction d'information dans des cas cliniques

Natalia Grabar (STL, CNRS, Université de Lille)

Cyril Grouin (LIMSI, CNRS, Université Paris Saclay)

Thierry Hamon (Université Paris 13, LIMSI, CNRS, Université Paris Saclay)

Vincent Claveau (IRISA, CNRS)

Introduction



Introduction

Analyse de cas cliniques rédigés en français

- Tâches : recherche d'information et extraction d'information
 - DEFT 2012, DEFT 2016 : identification de mots-clés
 - DEFT 2013 : appariement recette/ingrédients

Première fois sur des données cliniques en français !









Calendrier

- Communication : décembre 2018/avril 2019
 -  AIM, ARIA, EGC, Info-IC, LN, MadICS
 -  BioNLP, Corpora
- Accès aux données d'entraînement (licence) : 18 février
- Accès aux données de test (3 jours) : 9/15 mai

Licence


- Obligation de soumission de résultats et d'un article décrivant les méthodes
- Interdiction d'appartenir à l'un des trois laboratoires organisateurs
- Interdiction de redistribution des données, y compris dans le laboratoire, sans limite de temps

Participants

- 5 équipes académiques :
 - LGI2P/Mines Alès,  Nîmes
 - LIMICS/INRA (LAI),  Paris
 - LIPN/STIH,  Paris
 - TALN-LS2N,  Nantes
 - Université Assane Seck de Ziguinchor,  Sénégal
- 2 équipes industrielles :
 - EDF Lab,  Palaiseau
 - Qwant,  Paris
- 1 équipe mixte :
 - Synapse/IRIT,  Toulouse

Corpus

Corpus

- 718 cas cliniques  indexés (mots-clés) avec discussion [Grabar et al., 2019]
 - spécialités : cardiologie, gastro-entérologie, obstétrique, oncologie, pneumologie, urologie
 - période [2000–2018], parus :
 - en revue : *Pharmactuel*, *Progrès en Urologie*, *Revue des Maladies Respiratoires*, etc.
 - ou sur des sites spécialisés : *Association Française d’Urologie*, *Institut Marocain de l’Information Scientifique*, etc.
 - provenant d’un corpus annoté plus vaste [Grabar et al., 2018]
- Quelques publications rassemblent plusieurs cas cliniques
- Découpage des documents :
 - Cas clinique : autant de fichiers que de cas
 - Mots-clés : dupliqués pour chaque cas (même liste de mots-clés en plusieurs exemplaires)
 - Discussion : dupliquée pour chaque cas (même discussion en plusieurs exemplaires)

Corpus

Double annotation indépendante puis consensus, sur 4 catégories (informations démographiques et cliniques générales)

- **âge** : valeur + unité
- **genre** : indice textuel permettant d'inférer le genre parmi deux catégories (*femme, homme*)
- **issue** : phrase ou portion de phrase permettant d'inférer l'issue parmi cinq catégories (*guérison, amélioration, stable, détérioration, décès*)
- **origine** : phrase ou portion de phrase précisant l'origine de la consultation

Corpus

- **Guérison** : le problème clinique décrit est traité et la personne guérie
 - *Le recul était de deux ans sans récurrence locale ni incident notable*
- **Amélioration** : l'état clinique est amélioré sans conclure à une guérison
 - *Les suites ont été simples*
- **Stable** : état clinique stationnaire, choix impossible amélioration/détérioration
 - *La patiente présente toujours une constipation opiniâtre terminale, équilibrée sous traitement médical*
- **Détérioration** : dégradation de l'état clinique
 - *Un mois plus tard, le patient a été hospitalisé pour toxoplasmose cérébrale et pneumocytose pulmonaire, actuellement en cours de traitement*
- **Décès** : si le décès concerne directement le cas clinique décrit
 - *Le patient est décédé au 6ème mois après l'intervention*

Corpus

Annotation et normalisation (genre, issue) pour produire la référence de la tâche d'extraction d'informations

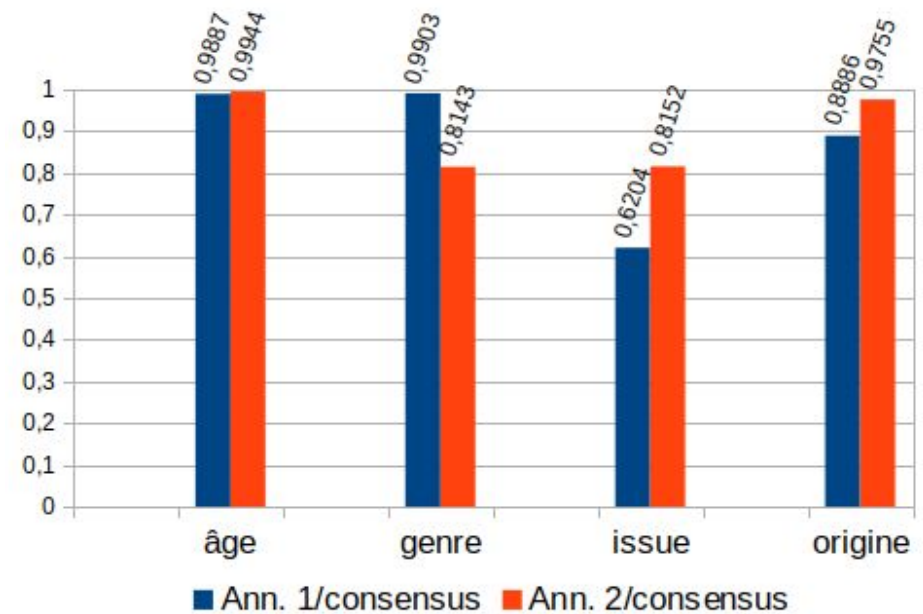
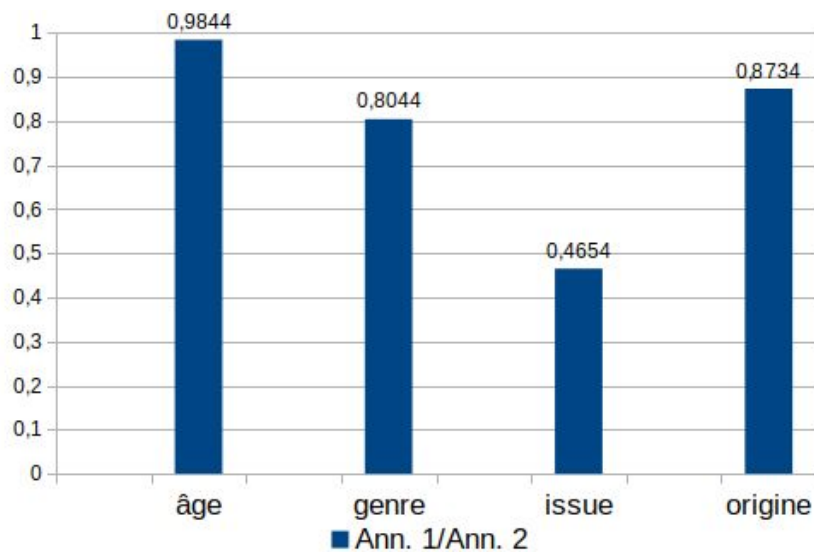
GEN [masculin]
Mr. H.J., âgé de âge 60 ans, ayant dans les antécédents des douleurs de la fosse iliaque droite avec hématurie épisodique, a été hospitalisé en origine urgence pour origine masse de la fosse iliaque droite fébrile avec pyurie.

issue [décès]
Le patient est issue [décès] décédé au 6ème mois après l'intervention.

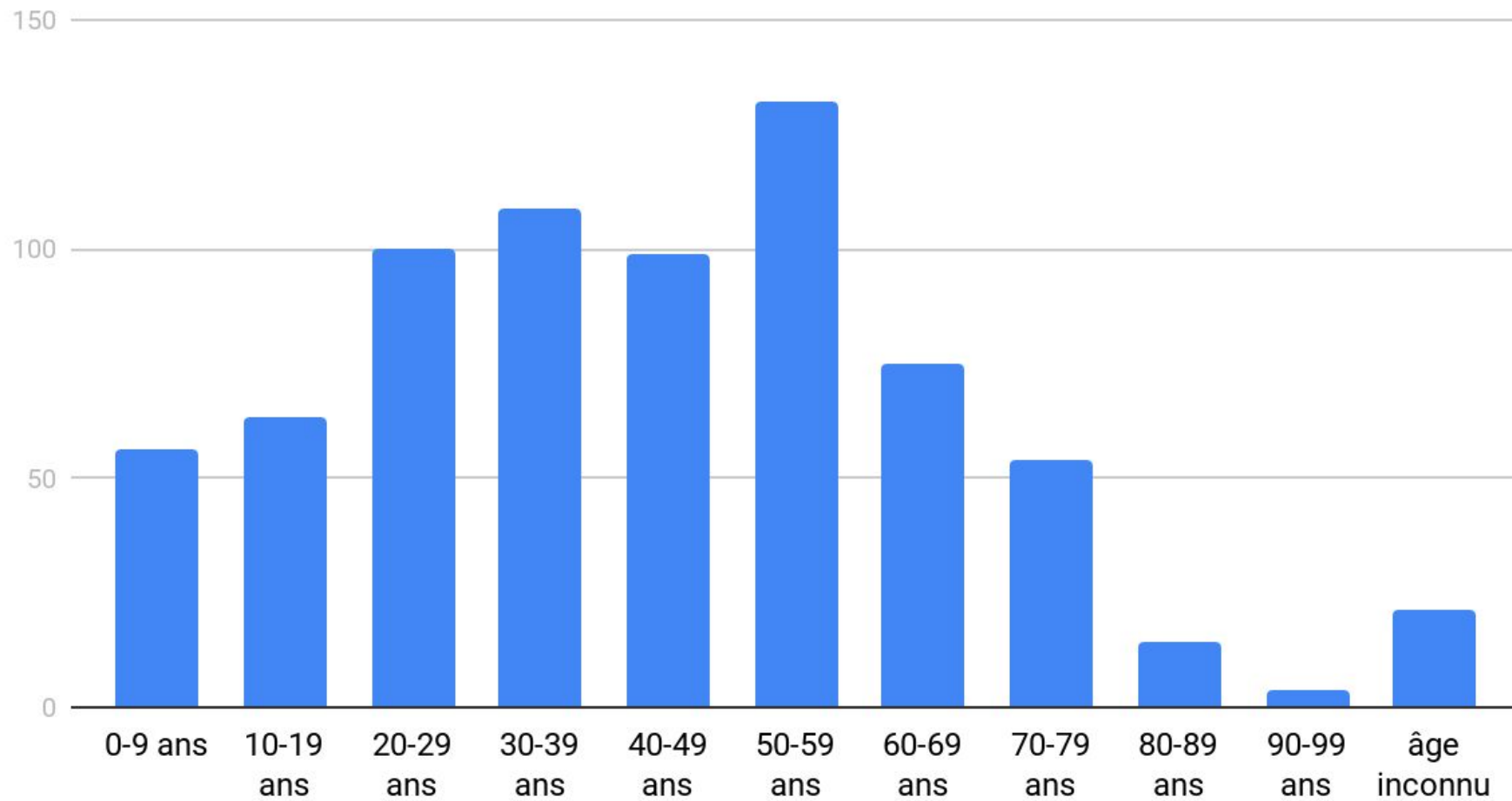
Corpus

Accords inter-annotateurs ainsi qu'entre annotateur et consensus (F-mesure)

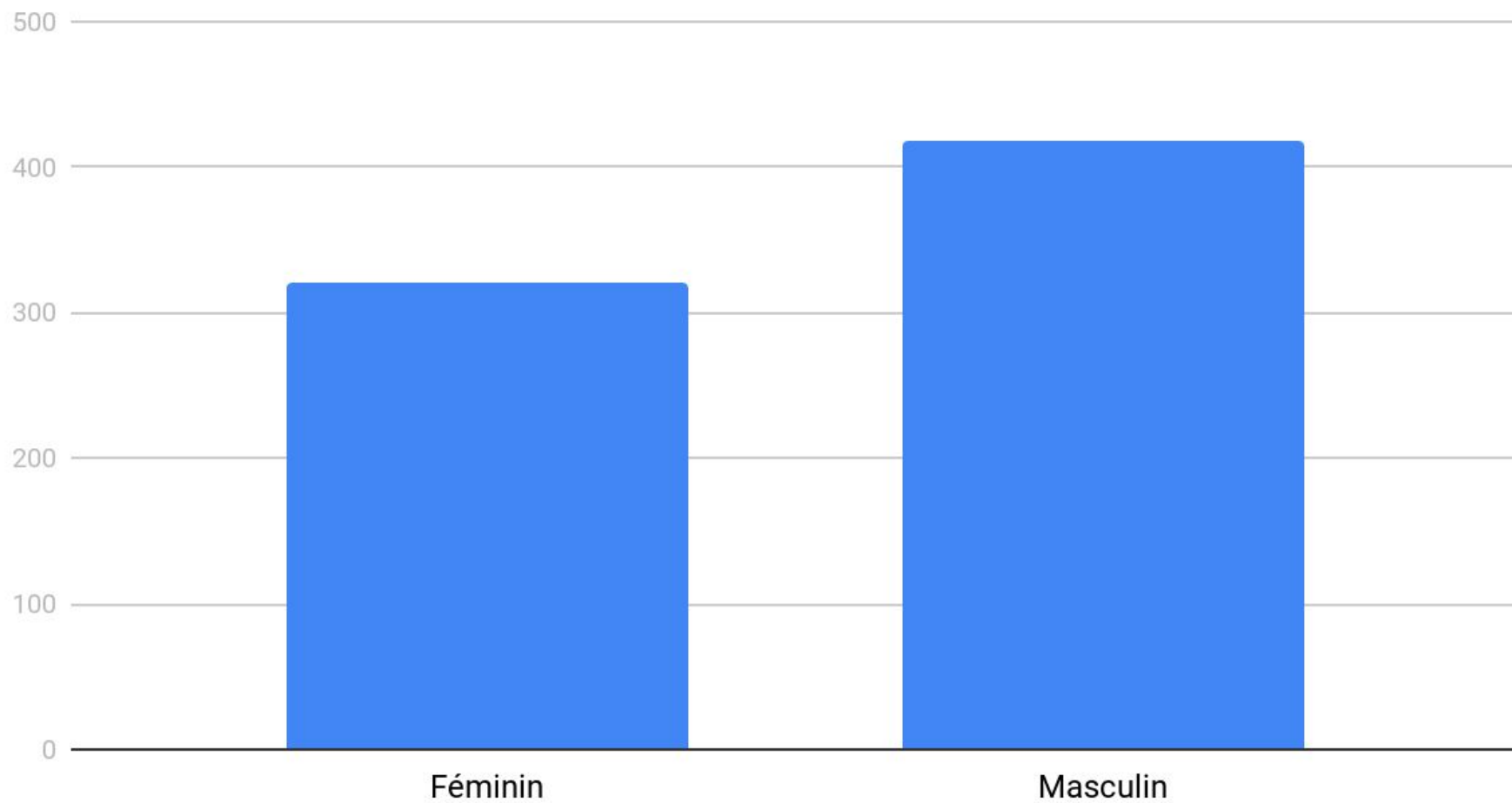
- Excellent accord (âge), très bon (origine, genre), faible (issue)
- Oublis, frontières, interprétation entre issues proches



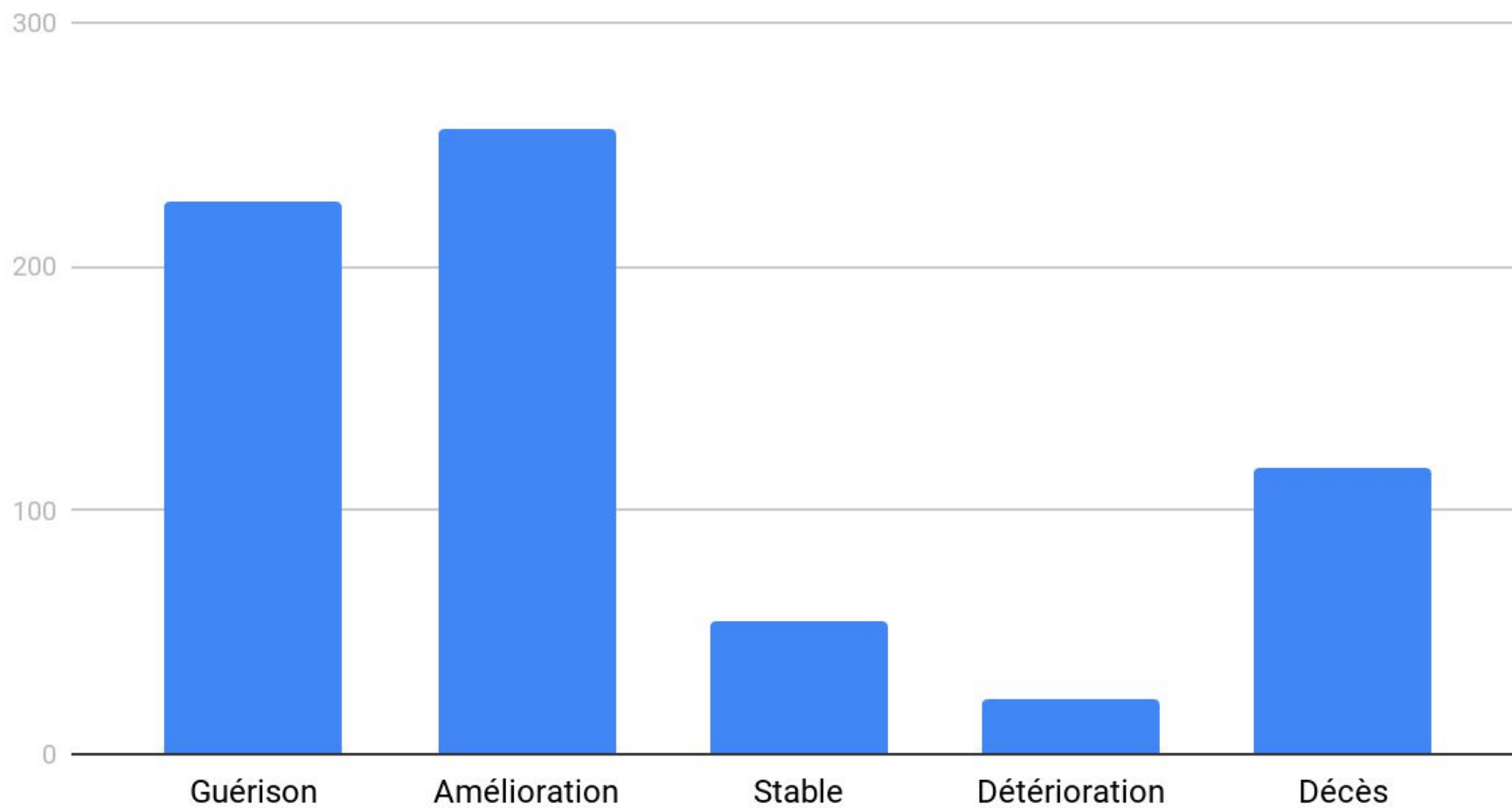
Distribution des tranches d'âges



Distribution des genres



Distribution des issues



Tâches

Tâche 1 : indexation de cas cliniques

- **Cas d'usage** : indexer/décrire un cas médical
- Entrée : cas clinique avec discussion correspondante, liste des mots-clés utilisés sur l'ensemble du corpus + nombre de mots-clés attendus
- Sortie : mots-clés pour chaque couple cas clinique/discussion, classés par ordre de pertinence décroissant
- Evaluation : MAP (mean average precision) + R-précision (secondaire)

Mot-clé	Cas clinique et discussion	Sous-corpus
agénésie	1136550700.txt 2300836250.txt	Entraînement
agénésie déférentielle	1139700160.txt 2354143280.txt	Test
agénésie rénale	Inutilisés dans la tâche d'indexation en 2019	
agénésie rénale unilatérale		

Tâche 2 : similarité sémantique cas / discussions

- **Cas d'usage** : trouver l'expertise pour un cas médical
- Entrée : cas cliniques + discussions correspondant
 - Remarque : une même discussion peut s'appliquer à plusieurs cas cliniques
- Sortie : appariement cas clinique/discussion
- Evaluation : précision
 - Remarque : les fichiers de discussion sont dédoublonnés. Il suffit qu'un fichier de la liste de discussion doublons soit trouvé pour une évaluation favorable

Tâche 3 : extraction d'information

- **Cas d'usage** : décrire finement le patient pour raisonnement, épidémiologie...
- **Âge** de la personne dont le cas est décrit, au moment du dernier élément clinique rapporté, normalisé sous la forme d'un entier
 - 0 pour un nourrisson de moins d'un an
 - 1 pour un enfant de moins de deux ans
 - 20 pour un patient d'une vingtaine d'années
- **Genre** de la personne (*féminin, masculin*)
- **Origine**, ou motif de la consultation, pour le dernier événement clinique ayant motivé la consultation
 - Pathologies, signes ou symptômes : *“une tuméfaction lombaire droite, fébrile avec frissons”*
 - Circonstances d'un accident : *“un AVP moto”, “pense avoir été violée”*
- **Issue** (*guérison, amélioration, stable, détérioration, décès*)

Résultats

Résultats

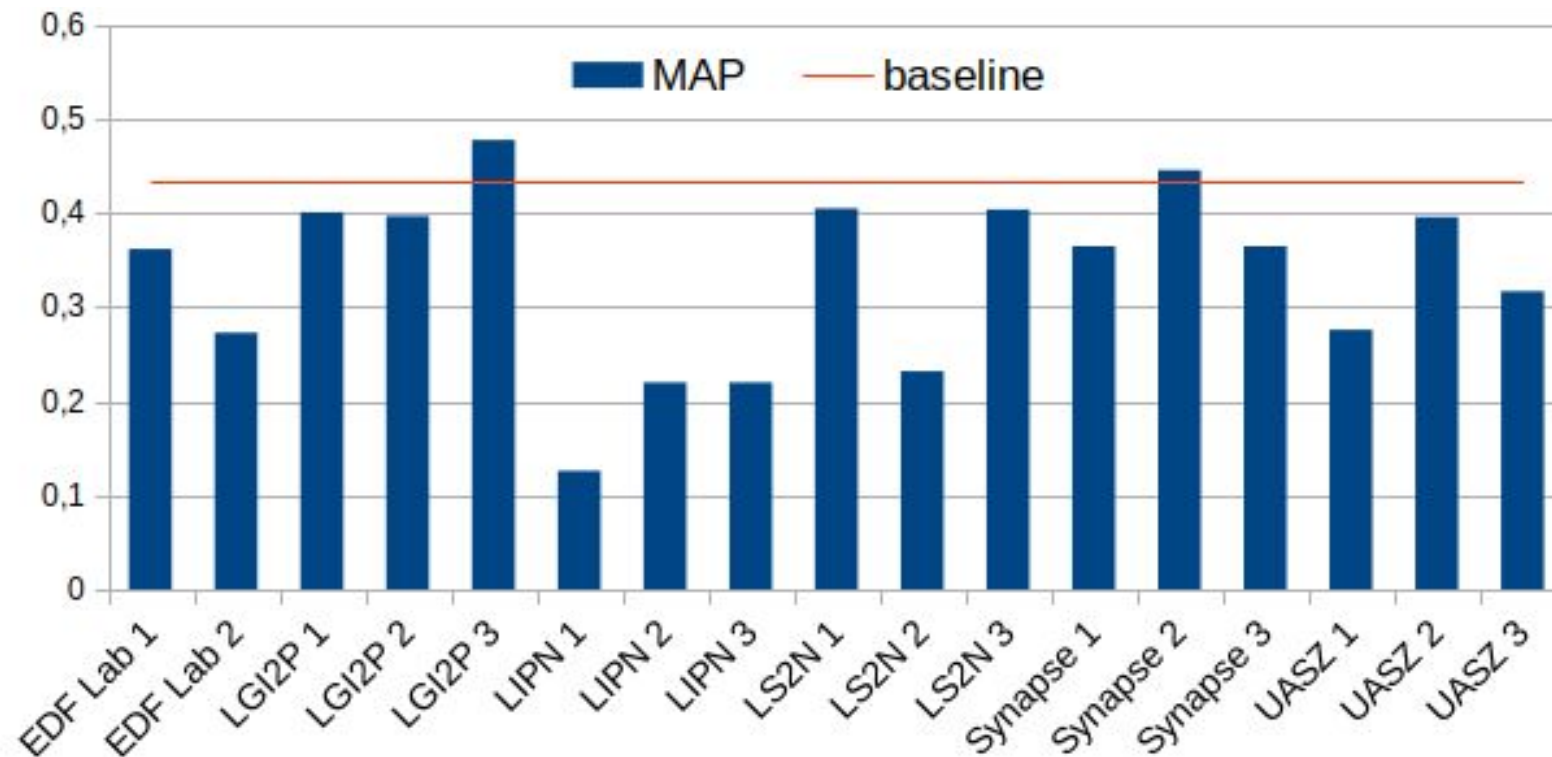
Baseline

- techniques simples/connues, ne nécessitant pas de données externes
- approches par règles ou apprentissage (non neuronal)

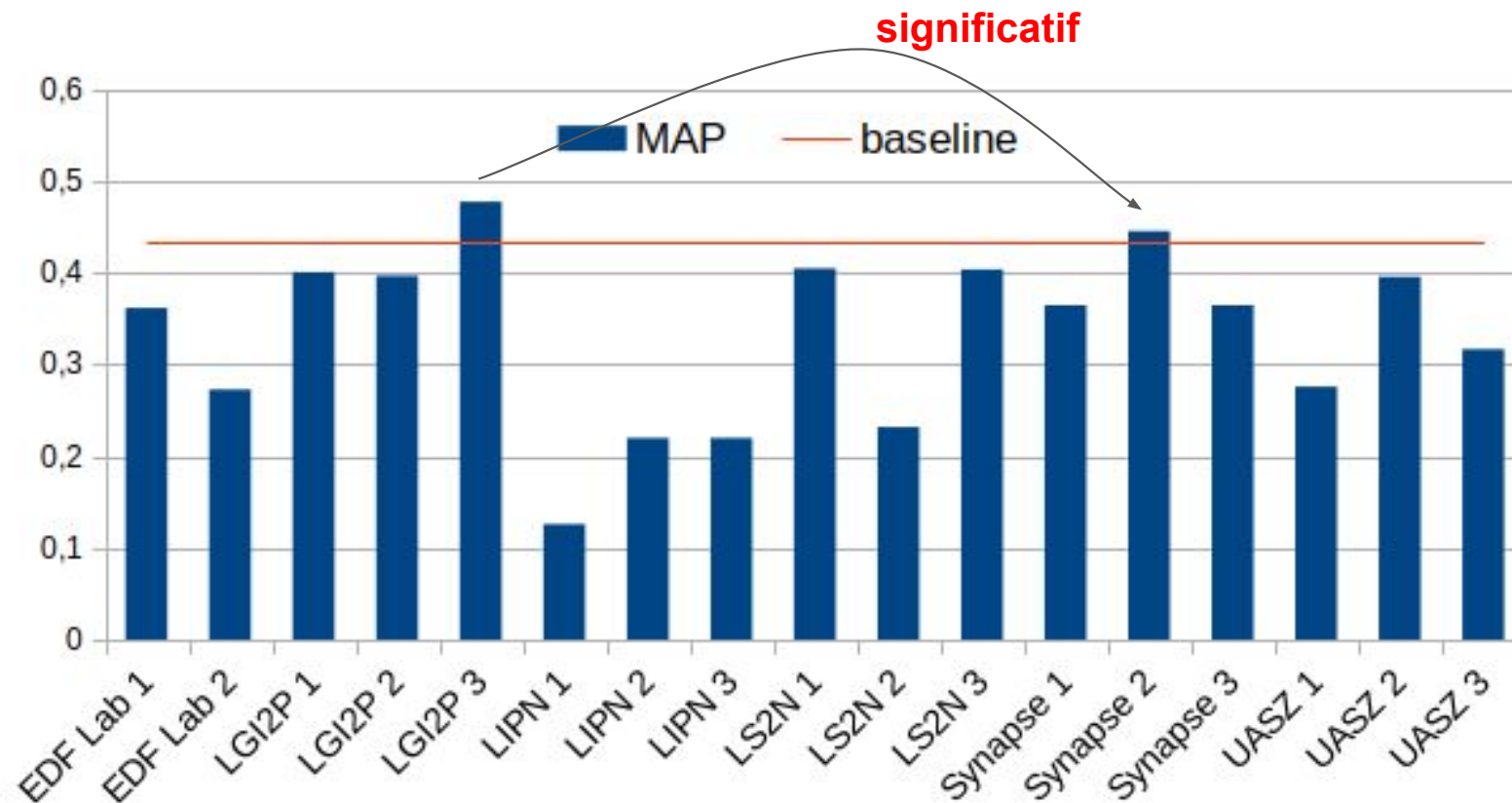
Significativité

- T-test pairé avec $p=0,05$

Tâche 1 : indexation des cas cliniques



Tâche 1 : indexation des cas cliniques



Commentaires tâche 1

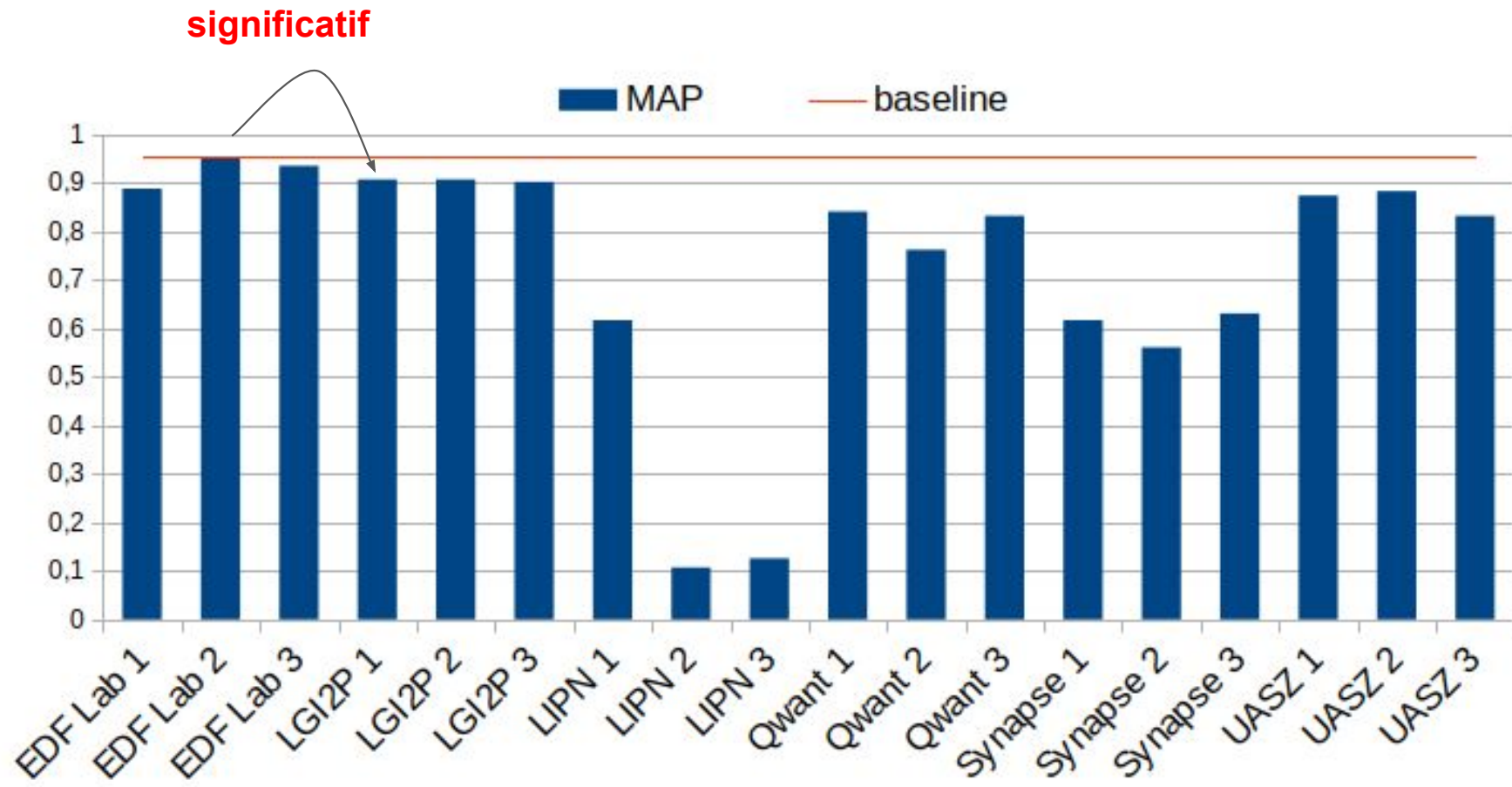
Significativité

- LGI2P > les autres
- Synapse ✗ baseline ; Synapse > LS2N

Approches

- approches RI classique, différentes pondérations
- plongements de mots
- classification (NB, gradient boosting...)

Tâche 2 : similarité sémantique cas/discussions



Commentaires tâche 2

Significativité

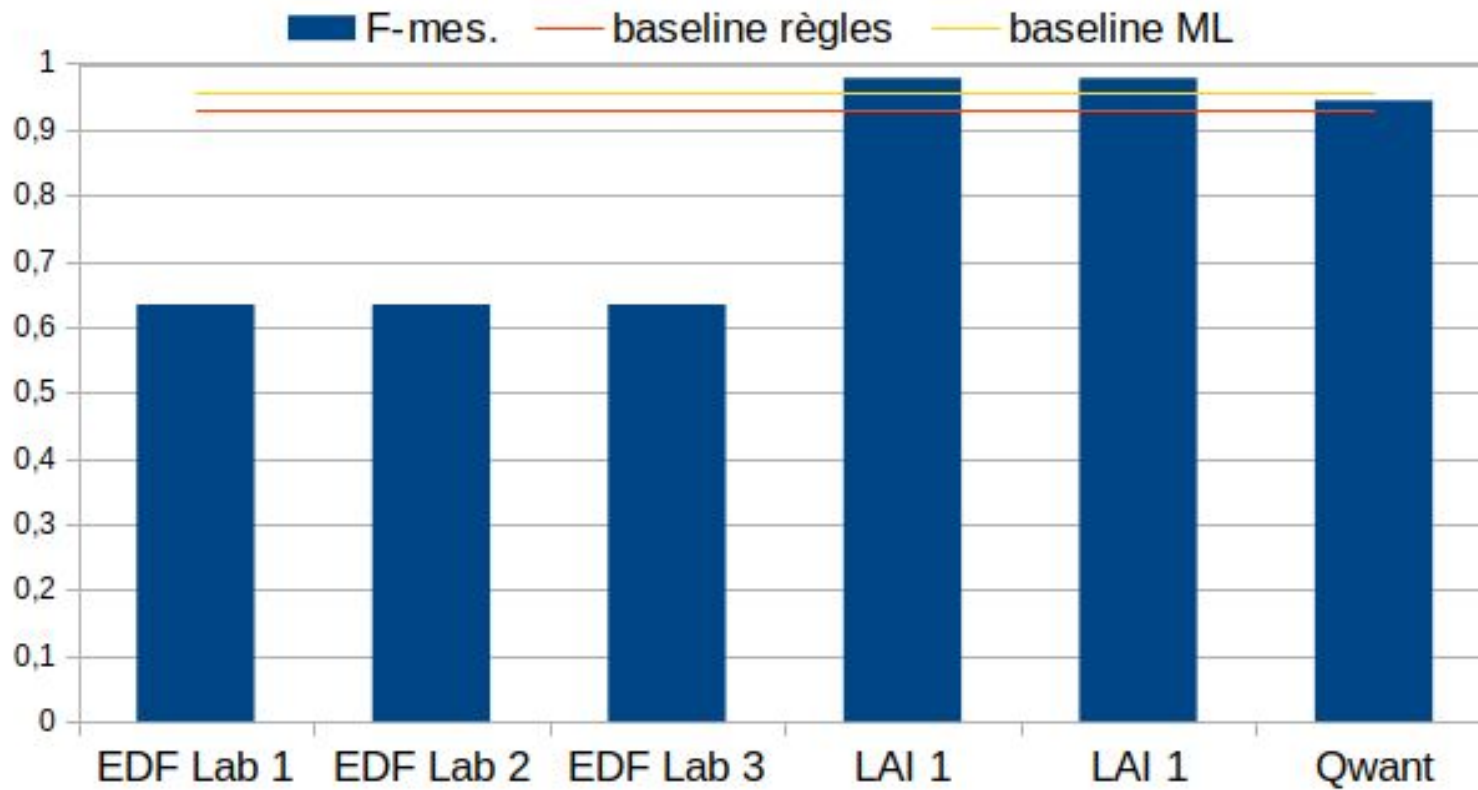
- résultats serrés mais EDF Lab run2 \succ les autres
- pas significatifs : EDF Lab (run 3) vs. LGI2P (run 1), LGI2P (run 1) vs. LGI2P (run 2), LGI2P (run 2) vs. LGI2P (run 3), LGI2P (run 3) vs. EDF Lab (run 1)

Approches

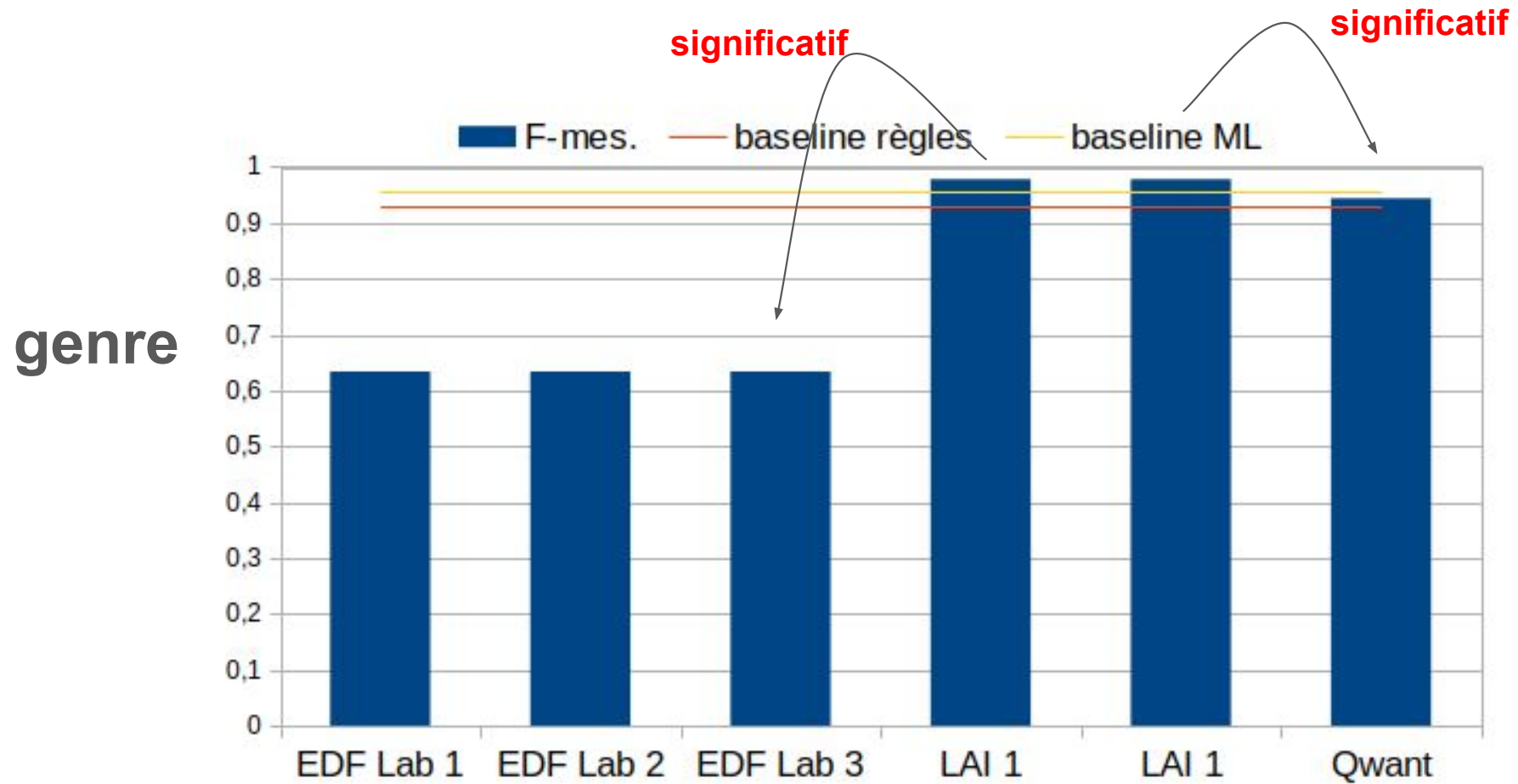
- représentations vectorielles (word2vec, LSI)
- modèles de langues
- CNN
- appariement optimal par algorithme hongrois

Tâche 3 : extraction d'informations

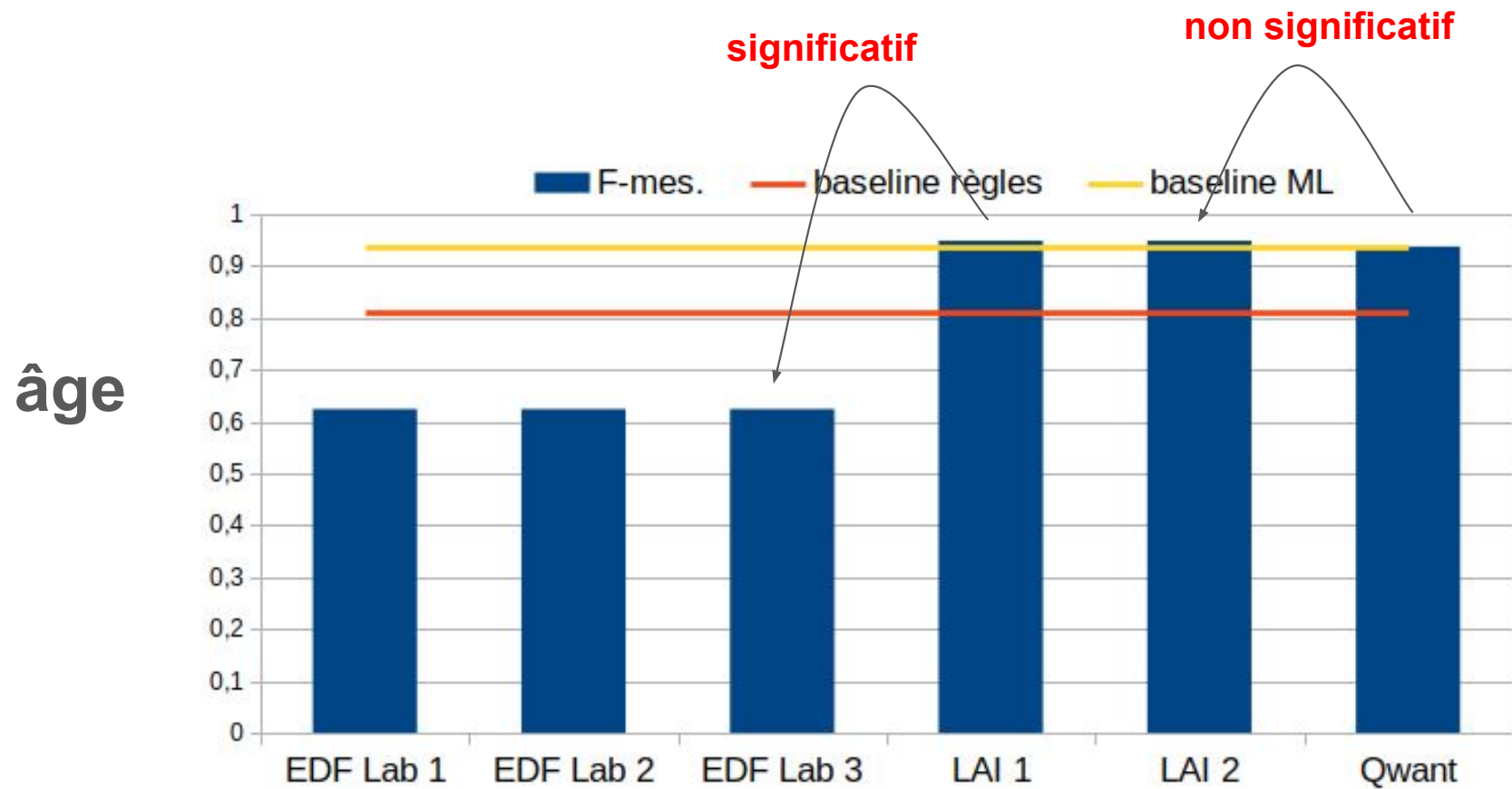
genre



Tâche 3 : extraction d'informations

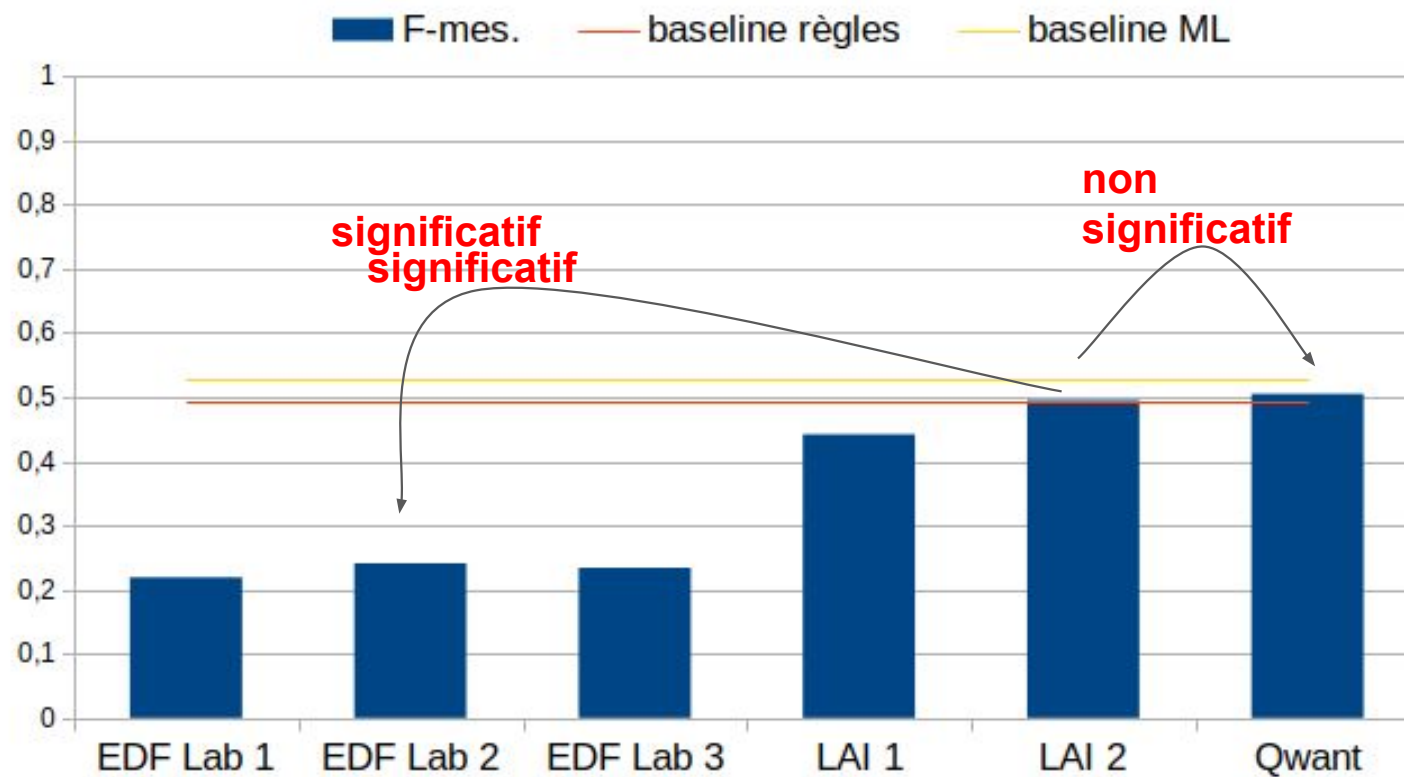


Tâche 3 : extraction d'informations



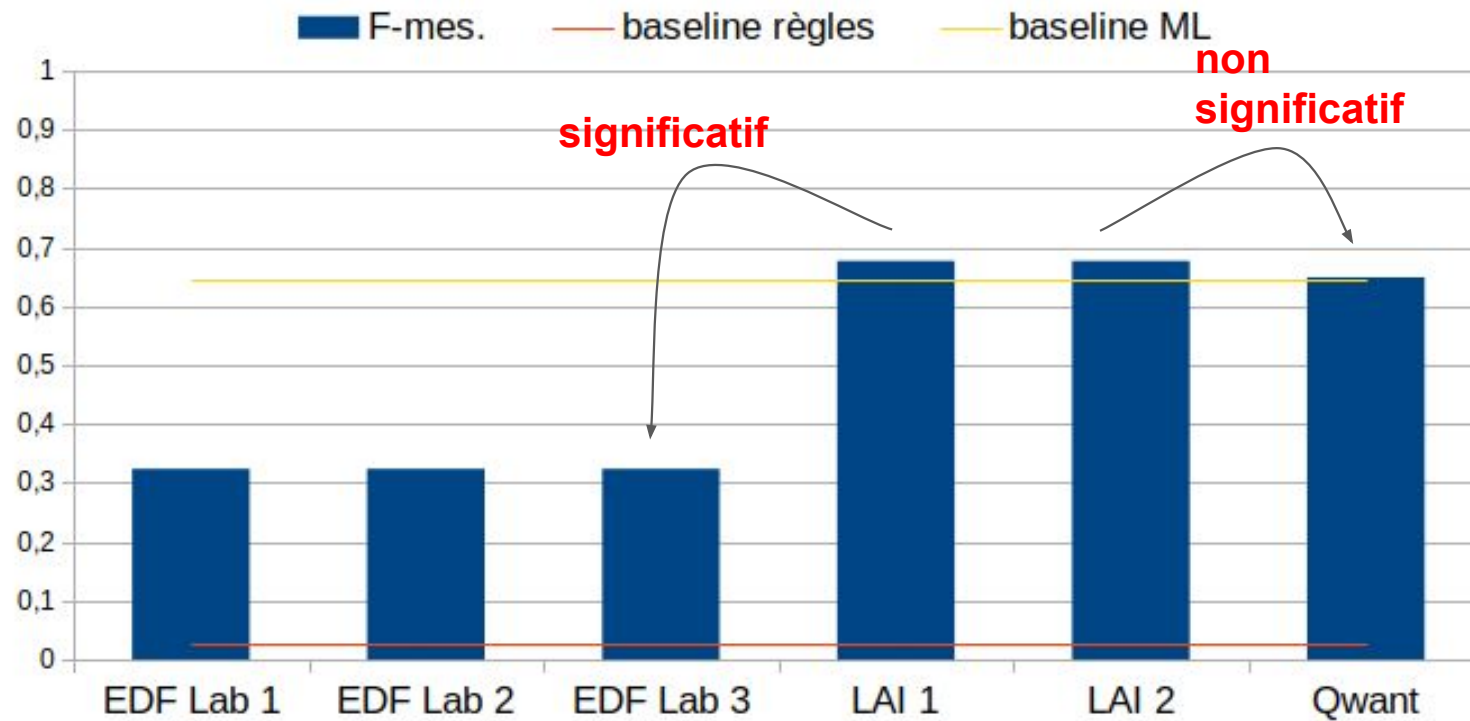
Tâche 3 : extraction d'informations

issue



Tâche 3 : extraction d'informations

origine



Commentaires tâche 3

Significativité

- toutes informations : LAI (run 2) et Qwant > EDF Lab (run 2)
- *genre* : LAI (run2) > Qwant
- *origine* (accuracy) : Qwant et LAI > baselines

Approches

- classification (genre, origine) vs. étiquetage (âge, issue)
- par règles ou lexiques (genre, origine)
- par apprentissage (CRF / RNN-CRF)

Conclusions

Participation

Succès en nombre de participants

Approches

- résultats assez serrés... mais des différences significatives
- peu de données pour les méthodes neuronales, mais des résultats intéressants
- quelques bugs dans certains runs

Futur

Suites directes

- analyse des cas d'erreurs
- complémentarité des approches, combinaison des approches

Cas cliniques

- développement du corpus
- annotation du corpus
- nouvelles tâches d'extraction d'information

Futur

DEFT 2020

- Mêmes tâches sur une autre partie du corpus : évolution des méthodes
- Catégories EI proches des enjeux cliniques [Grouin et al., LOUHI 2019]

, les diagnostics évoqués étaient une pathologie tumeur primitive du ANAT foie ou une PATHO tumeur de la anatomie surrénale. I

it d'une PATHO tumeur de la anatomie veine cave inférieure sus-rénale. Une TRT exérèse de la PATHO tumeur et du segment
prothèse synthétique est mise en place assurant la continuité du retour veineux. Les suites imm

es SOSY œdèmes des anatomie membres inférieurs. L'examen échodoppler montrait une pathologie thrombose étendue depuis la

orait cliniquement après administration de l'substance héparine relayée par une classe anti-vitamine K. L'examen

Futur

Futures éditions de DeFT

Souhaits de tâches ?

Idées de corpus ?

Remerciements

- Participants : intérêt pour les tâches proposées, méthodes développées
- Projet CLEAR (ANR-17-CE19-0016-01)
- LabEx COMIN LABS (ANR-10-LABX-07-01)