

# Classification d'Alzheimer à partir de paramètres acoustiques et prosodique avec de l'apprentissage automatique

Eric Jordan\* Jalal Al-Tamimi\*

\*Université Paris Cité, Laboratoire de linguistique formelle (LLF), F-75013 Paris France



## Objectifs

- Quantifier les différences entre participants atteints d'Alzheimer (Alz) et contrôles neurotypiques (CN)
- Comparer la performance des mesures automatiques de Prosogram et d'OpenSMILE
- Utiliser des algorithmes interprétables et explicables - pas de boîte noire!

## Introduction

### Maladie d'Alzheimer

- La maladie d'Alzheimer (MA) est une maladie neurodégénérative qui engendre un déclin des capacités cognitives du patient
- MA impacte aussi sur la parole, sur les aspects temporels (plus de pauses, débit réduit), fréquentiels et la qualité vocale (plus de souffle) (Ivanova et al., 2022).

## Données

### L'ensemble de données ADRess

- Enregistrements de parole spontanée en anglais issus d'une tâche de description d'image (dite "Cookie Theft")
- 156 participants (86 F, 70 H) âgés entre 50 et 80 ans, ils sont à 50% atteint d'Alz et 50% CN
- Les participants sont à 50% atteint d'Alz et 50% CN
- Audio brut (plus que 3 heures) et segmenté (activité vocale  $\geq 65$ dB jusqu'à 10 seconds)
- Pré-répartie en base d'apprentissage et de test (108/48, équilibré en termes d'âge et de genre)
- Créée initialement pour l'ADReSS Challenge, un défi qui a eu lieu à l'Interspeech 2020. Des résultats de références sont fournis avec un taux de classification correcte (TCC) maximal de 62,5% sur des paramètres acoustiques.

## Les ensembles de mesures

- Ensemble de données ADRess - audio brut et segmenté (activité vocale  $\geq 10$ s)
- 2 Ensembles de mesures OpenSMILE, ComParE et GeMAPS. Chacun avec 2 variants: Profile spectral moyen (Functionals), mesures acoustiques par trame (LLDs)
- GeMAPS contient des mesures spectrales, temporelles, fréquentielles, d'amplitude et de qualité vocale. ComParE comprend des mesures spectrales, cepstrales, temporelles, fréquentielles, d'amplitude et de qualité vocale. Les 2 ensembles Functionals rajoutent des mesures temporelles.

- Prosogram a été utilisé afin d'automatiser la segmentation de l'audio ainsi que l'extraction de mesures spécifique à Prosogram
- L'ensemble de mesures de Prosogram couvre des mesures fréquentielles et temporelle

Feature Set	LLDs	Functionals
ComParE	65	6373
GeMAPS	18	62
Prosogram	-	43

Table 1. Nombre de mesures

## Méthodologie

### Extraction des mesures

- Prosogram a été utilisé afin de segmenter l'audio brut des enregistrements ADRess
- Via Python les mesures OpenSMILE ont été extrait des segments ADRess et des segments provenant de Prosogram (8 ensembles de paramètres)
- Segmentation avec Prosogram a engendré une perte de données car certaines segmentations étaient trop courte pour la taille de fenêtre utilisée par OpenSMILE
- L'ensemble de mesures de Prosogram ont été extraits via Praat (1 ensemble de paramètres)
- Les mesures OpenSMILE extraites sur les segments normalisés d'ADReSS ont été moyennées afin d'avoir une observation par participant. Ceci permet de les comparer directement aux mesures de Prosogram, pour un nombre total d'ensembles de mesures de 13

## Algorithmes d'apprentissage automatique

- Après normalisation (cote Z) et pré-traitement (suppression de mesures colinéaires et élimination récursive des caractéristiques (ERC))
- 7 algorithmes d'apprentissage automatique ont été choisis selon soit leur familiarité et utilisation commune (Régression Logistique, Machine à Vecteurs de Support (MVS)) ou leur applicabilité au contexte (Forêt d'arbres de décision, version ExtraTrees (ET))
- Performance quantifiée avec le TCC, le score F1 (précision/recall), AUC, sensibilité et spécificité
- Évaluation de prédictions

## Résultats

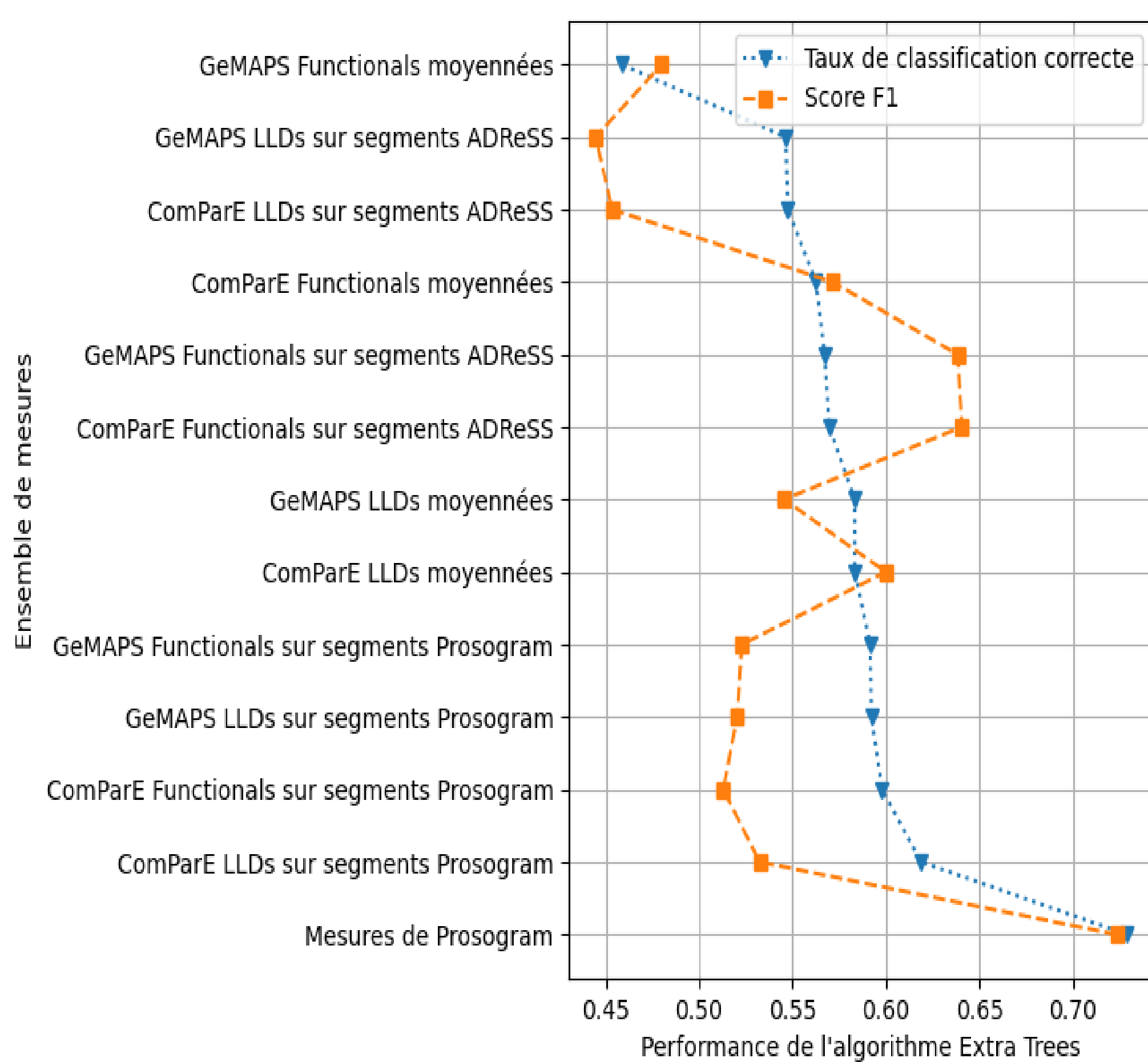


Figure 1. Résultats de l'algorithm Extra Trees sur tous les ensembles de mesures

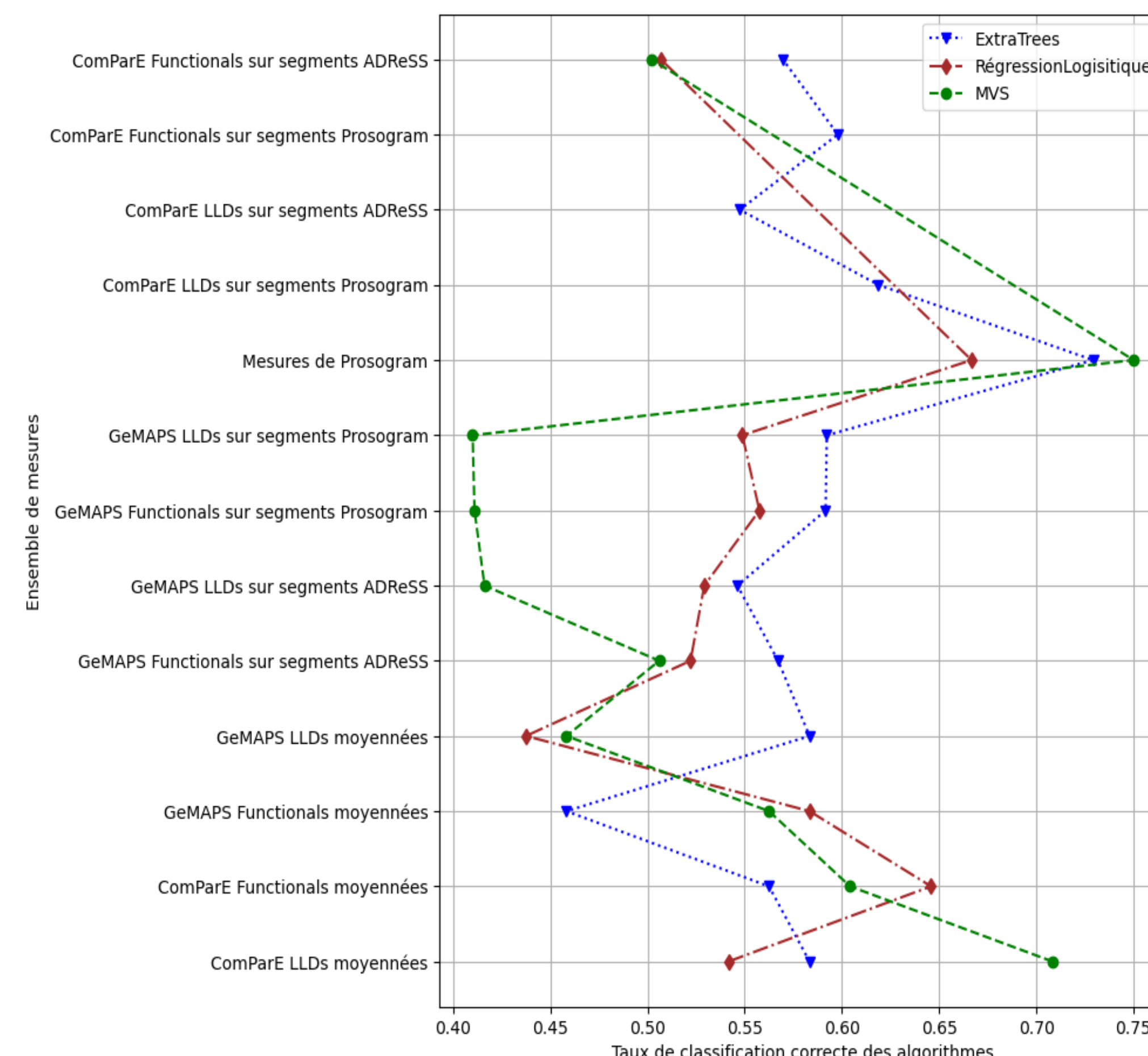


Figure 2. TCC des algorithmes les plus performants

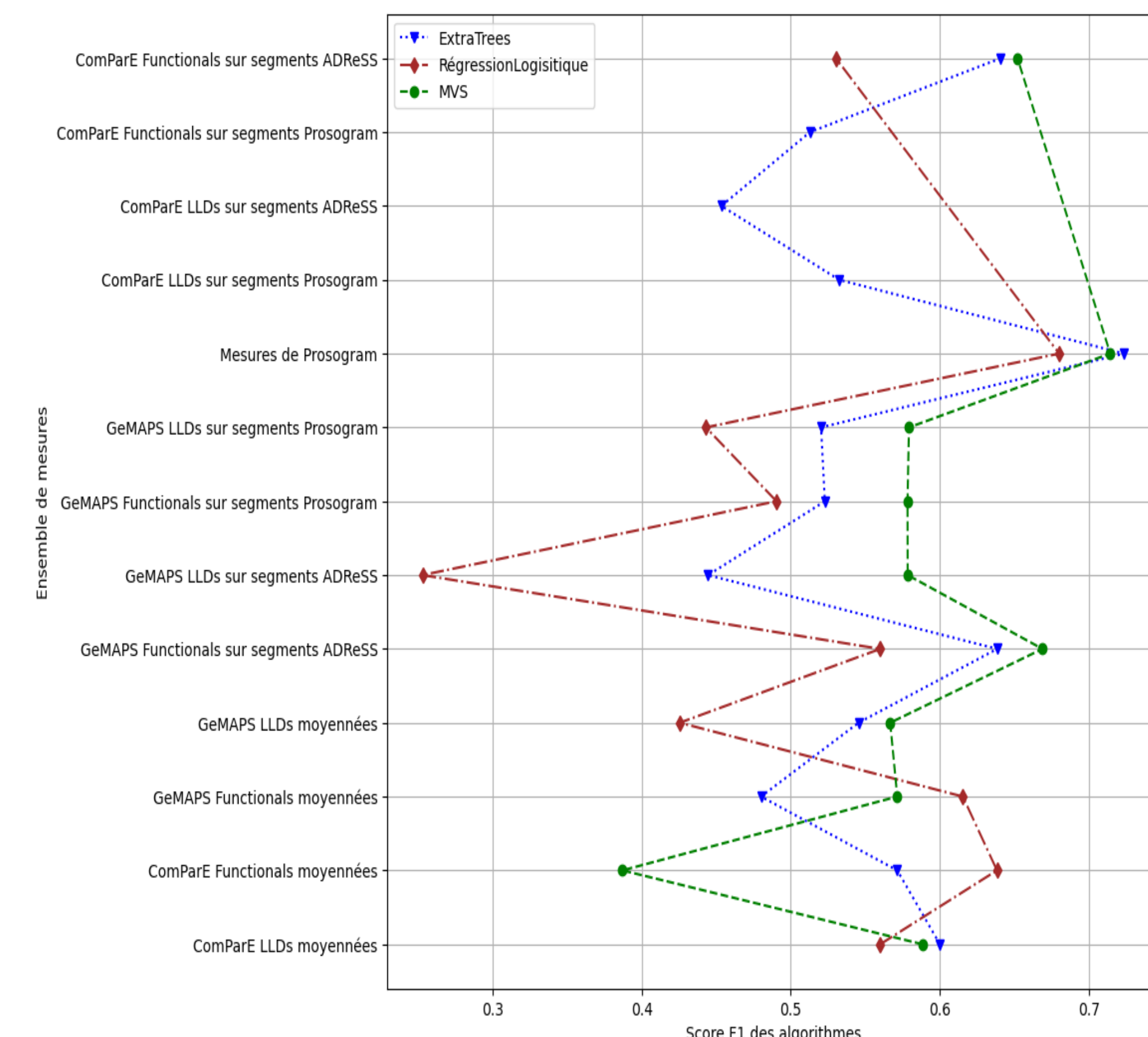


Figure 3. Score f1 score des algorithmes les plus performants

## Analyse des Résultats

- Meilleure performance globale atteinte avec l'algorithm ExtraTrees avec un TCC moyen de 58%, la meilleure performance absolue était un TCC 75% (MVS sur les mesures de Prosogram)
- Les meilleures performances ont été réalisées par les données avec peu d'observations (Prosogram et OpenSMILE moyenné)
- Sur l'audio normalisé, le TCC des ensembles fonctionnels était plus élevé par rapport aux ensembles de LLDs sur les mêmes données

## Conclusion

- L'ensemble de mesures de Prosogram ainsi que la segmentation automatique via Prosogram ont fait preuve de résultats prometteurs avec un TCC plus élevé que celui des résultats de référence ADRess
- L'algorithm ET était un bon candidat pour l'ERC à cause de sa bonne performance globale
- Le TCC plus élevé des ensembles fonctionnels par rapport aux ensembles de LLDs ainsi que la forte performance des mesures de Prosogram (avec un grand nombre de mesures temporelles) indiquent, comme d'autres études le suggèrent, que les mesures temporelles sont les plus importantes pour distinguer entre une population Alz et CN