

DÉTECTION PRÉCOCE DE LA MALADIE D'ALZHEIMER PAR ETUDE DE L'EXPRESSION ORALE RETRANSCRITE À L'ÉCRIT

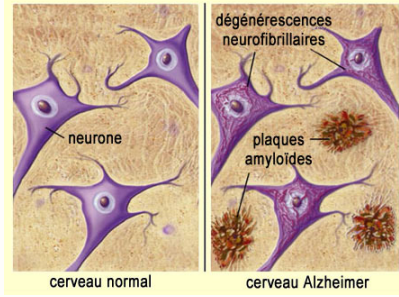
Gabriel Chabredier, Valentine Gobert, Mounir Lahlouh, Hela Marouane, Mohammed Ibn Kheder

SystemX

ESME Research Lab, Paris, France

Introduction : Alzheimer, la maladie

- Plaques amyloïdes (protéine bêta-amyloïde)
- Enchevêtrements neurofibrillaires (protéine Tau)



Conséquence: Mort Cellulaire

Méthodes de diagnostic:

- IRM
- Scanner
- Biomarqueurs

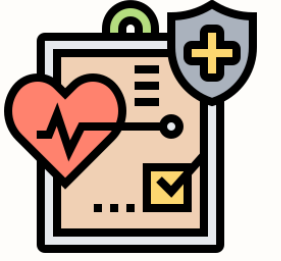
Enjeux de la détection précoce de la démence :



- 1) Maximiser l'efficacité des traitements
- 2) Prise en charge de la cellule familiale
- 3) Accessibilité à un diagnostic

Premiers signes:

- Mutisme
- Apraxie
- Agnosie



Etat de l'art

Référence	Dataset	Tâche	DL Network	Résultats
Pappagari et al., 2021	ADReSSO Dataset	AD-HC Classification	SpeechBrain ASR embedding + X-vector + BERT + LR fusion [1]	Précision: 84.51%
Qiao et al., 2021	ADReSSO Dataset	AD-HC Classification	Google Cloud ASR + BERT + Acoustic features sets + LR + Average late-fusion strategy [2]	Précision: 84.51%
Chen et al., 2021	ADReSSO Dataset	AD-HC Classification	2 LRs using complexity & disfluency features + BERT + ERNIE models [3]	Précision: 81.69%

Données

Challenge ADReSSO



DementiaBank Dataset

166 audios d'entraînement

- 79 patients malades
- 87 patients non-malades

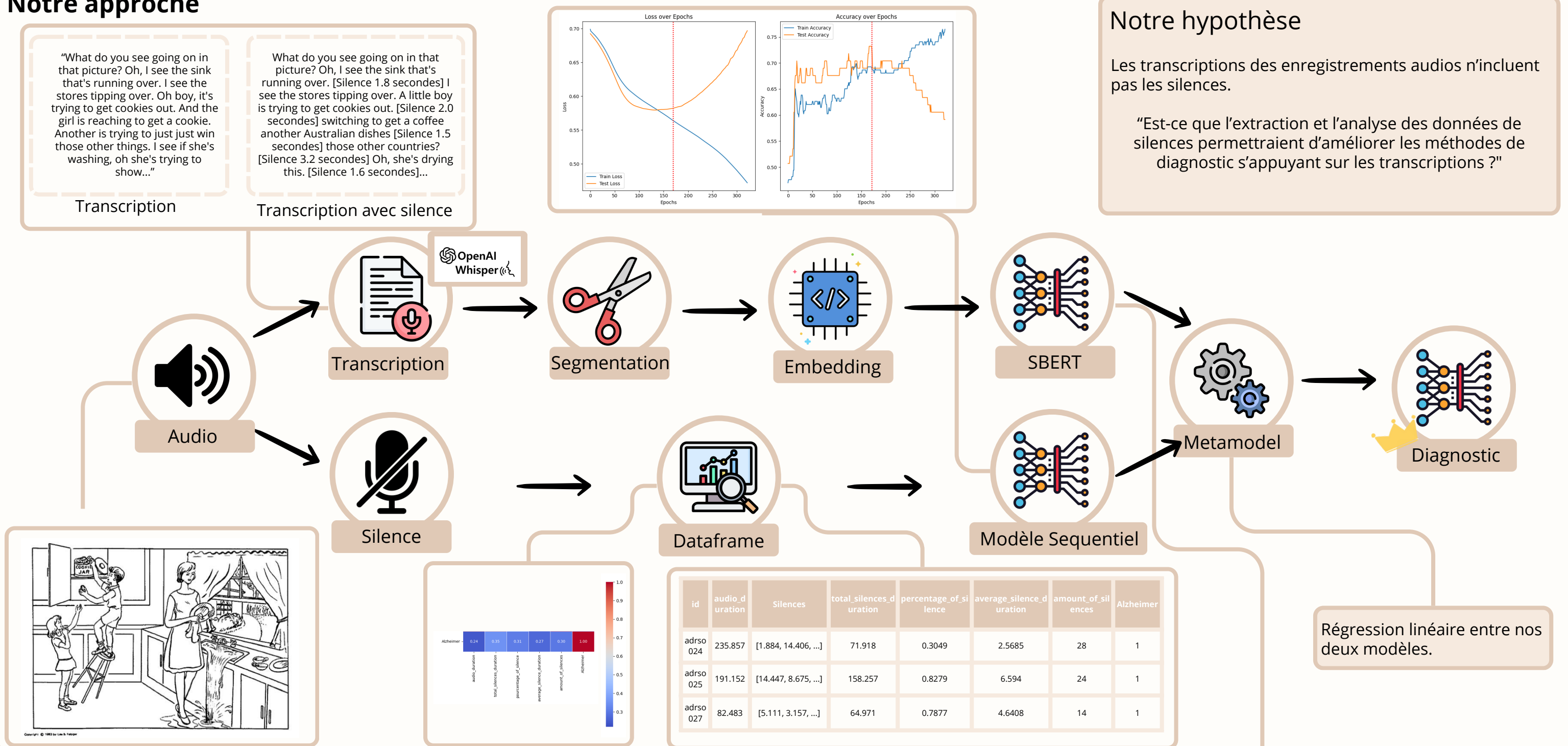
71 audios de test

- 35 patients malades
- 36 patients non-malades

Problématiques :

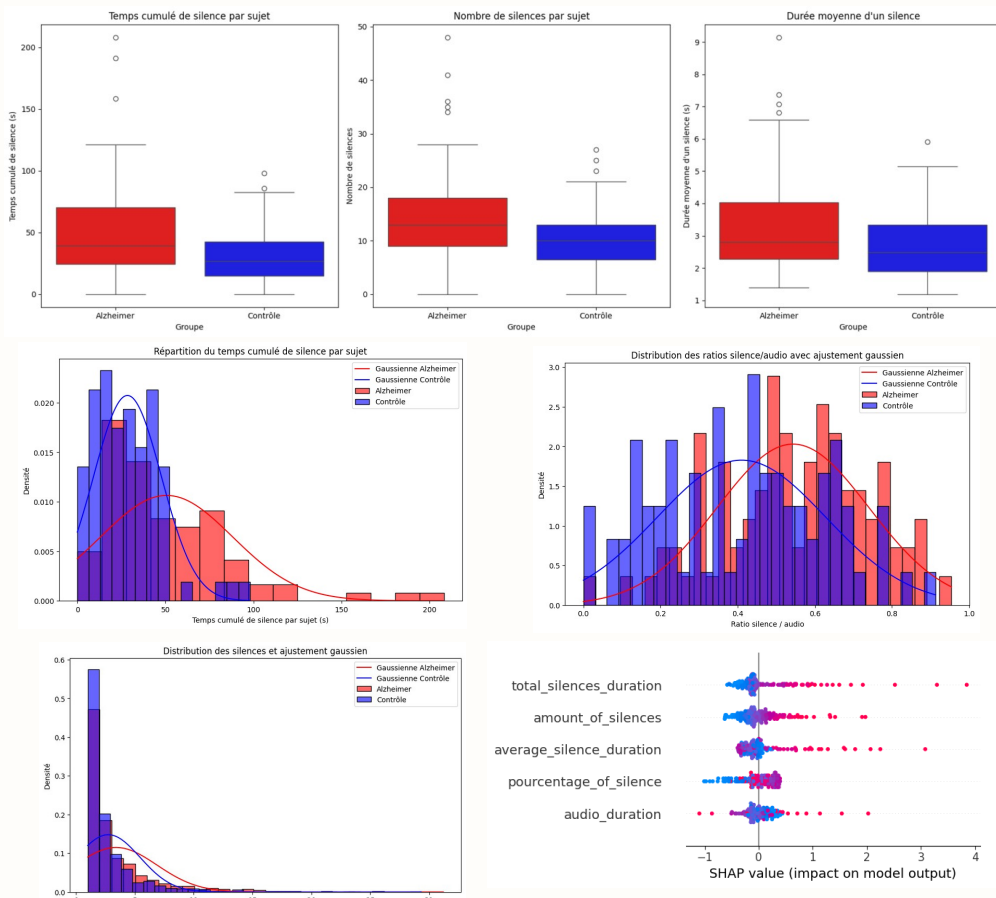
- Faible quantité de données
- Augmentation de données impossible
- Données Médicales Sensibles

Notre approche



Résultats

Pertinence des données de silence



Métriques du Metamodel

Données d'Entraînement

Metric	Value
Mean Squared Error	0.0775

Données de Test

Metric	Value
7 Mean Squared Error	0.1362

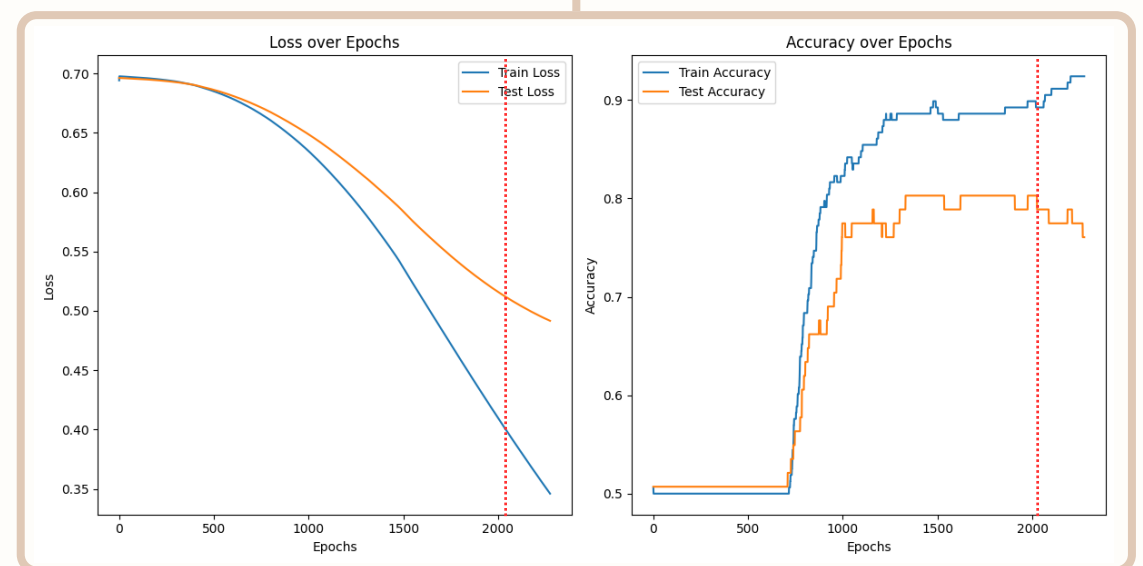
Après seuillage à 0.5

Données d'Entraînement

Metric	Value
0 Mean Squared Error	0.0775
1 AUC-ROC	0.9805
2 Accuracy	0.9367
3 True Positives	72
4 False Positives	3
5 False Negatives	7
6 True Negatives	76

Données de Test

Metric	Value
7 Mean Squared Error	0.1362
8 AUC-ROC	0.9119
9 Accuracy	0.8451
10 True Positives	30
11 False Positives	6
12 False Negatives	5
13 True Negatives	30



Conclusion

- Obtention de la même valeur de précision que les deux meilleures équipes de chercheurs qui ont participé au challenge ADReSSO qui était de 84.51%.
- Hypothèse prouvée, l'étude des silences permet l'améliorer des méthodes de diagnostic s'appuyant sur les transcriptions

Sources

- [1] Pappagari, R., Cho, J., Joshi, S., Moro-Velázquez, L., Żelasko, P., Villalba, J., & Dehak, N., 2021. Automatic Detection and Assessment of Alzheimer Disease Using Speech and Language Technologies in Low-Resource Scenarios. Interspeech 2021. 3825-3829. 10.21437/Interspeech.2021-1850.
- [2] Chen, J., Ye, J., Tang, F., & Zhou, J., 2021. Automatic detection of alzheimer's disease using spontaneous speech only. Interspeech 2021. 3830-3834. 10.21437/Interspeech.2021-2002.
- [3] Qiao, Y., Yin, X., Wiechmann, D., & Kerz, E., 2021. Alzheimer's Disease Detection from Spontaneous Speech Through Combining Linguistic Complexity and (Dis)Fluency Features with Pretrained Language Models. Interspeech 2021. 3805-3809. <https://doi.org/10.48550/arXiv.2106.08689>.