

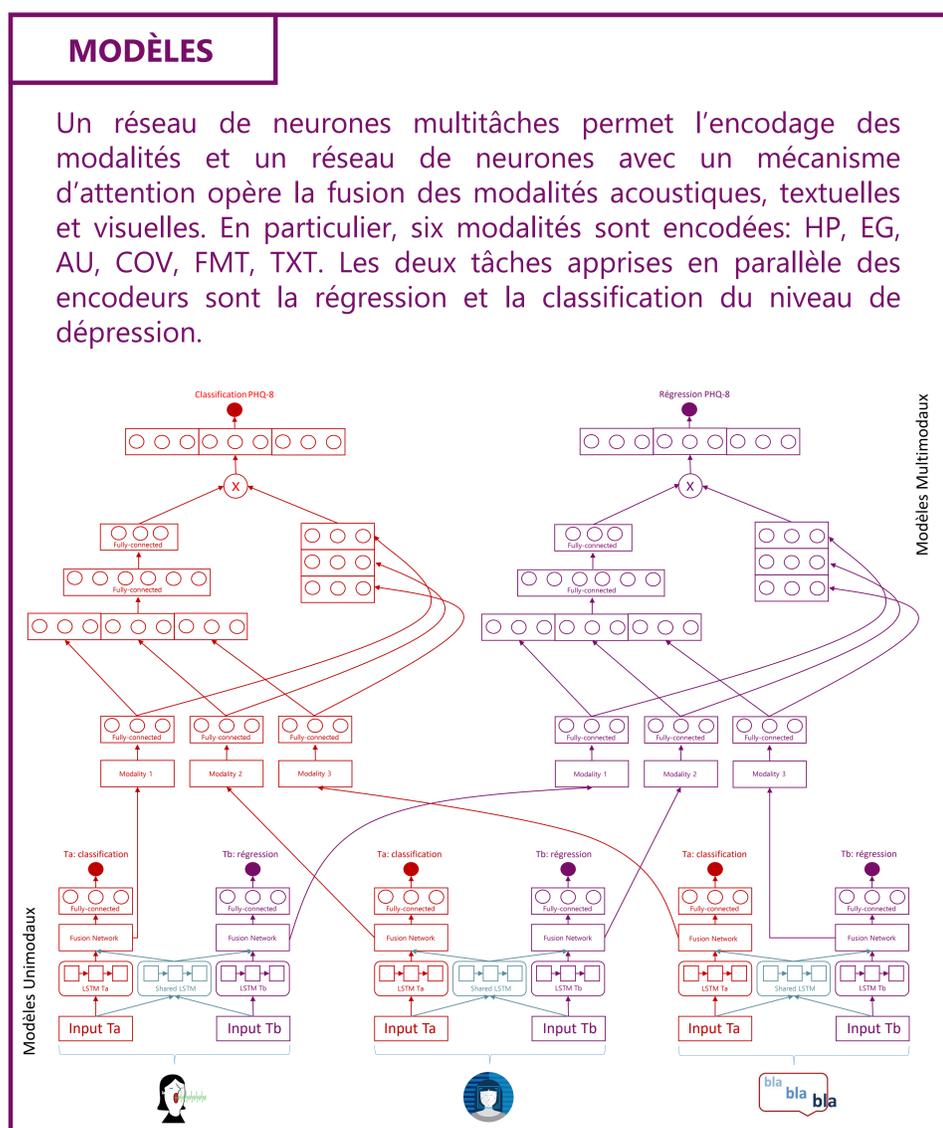
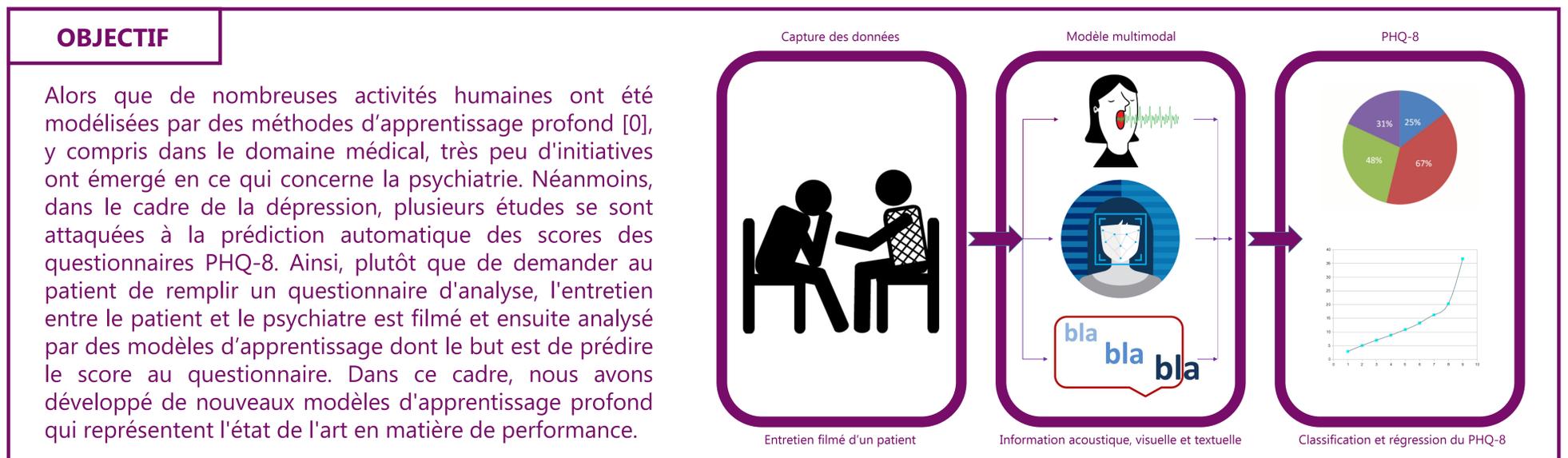
# Prédire les Évaluations PHQ-8 par Méthodes d'Intelligence Artificielle

Gaël DIAS<sup>1</sup>, Mohammed HASANUZZAMAN<sup>2</sup>, Sriparna SAHA<sup>3</sup>, Arbaaz QURESHI<sup>1,3</sup>

<sup>1</sup> Normandie Univ., UNICAEN, ENSICAEN, CNRS, GREYC, Caen, France. ✉ gael.dias@unicaen.fr

<sup>2</sup> ADAPT Center, Dublin, Ireland. ✉ mohammed.hasanuzzaman@adaptcentre.ie

<sup>3</sup> Indian Institute of Technology, Patna, India. ✉ sriparna@iitp.ac.in, arbaaz.queshi29@gmail.com



### RÉSULTATS

Les modèles ont été entraînés et testés sur un ensemble de 138 patients américains souffrant de dépression [1]. Deux types d'expériences ont été menées. Les modèles proposés doivent être capable de prédire le score du PHQ-8 sur une échelle continue de 0 à 24 (régression), et sur une échelle discrète de 0 (pas dépressif) à 5 (dépressif sévère) - classification. Les modèles développés dans cette étude [2] présentent des résultats supérieurs à l'état de l'art proposé par [3].

Network	RMSE	MAE	Acc (%)
ST. DLR. HP	6.89	5.67	-
ST. DLC. HP	-	-	54.54
FS. MT. HP	6.75	<b>5.48</b>	<b>60.60</b>
SP. MT. HP	<b>6.65</b>	5.53	54.54
ST. DLR. EG	6.67	4.72	-
ST. DLC. EG	-	-	54.54
FS. MT. EG	<b>6.50</b>	<b>4.60</b>	<b>57.57</b>
SP. MT. EG	6.59	5.16	57.57
ST. DLR. AU	6.49	5.55	-
ST. DLC. AU	-	-	54.54
FS. MT. AU	<b>6.28</b>	<b>5.03</b>	54.54
SP. MT. AU	6.46	5.42	<b>57.57</b>
ST. DLR. COV	6.64	5.72	-
ST. DLC. COV	-	-	51.51
FS. MT. COV	<b>6.55</b>	<b>5.67</b>	<b>54.54</b>
SP. MT. COV	6.59	5.71	54.54
ST. DLR. FMT	6.91	5.89	-
ST. DLC. FMT	-	-	51.51
FS. MT. FMT	6.72	<b>5.77</b>	<b>54.54</b>
SP. MT. FMT	<b>6.69</b>	5.79	51.51
ST. DLR. TXT	4.90	3.99	-
ST. DLC. TXT	-	-	60.60
FS. MT. TXT	4.96	3.90	<b>66.66</b>
SP. MT. TXT	<b>4.70</b>	<b>3.81</b>	60.61
ST. DLR. CombAtt	4.42	3.46	-
MT. DLR. CombAtt	<b>4.24</b>	<b>3.29</b>	-
ST. DLC. CombAtt	-	-	57.57
MT. DLC. CombAtt	-	-	<b>60.61</b>
VFSC <sub>sem</sub>	4.46	3.34	-
AW <sub>bhv</sub>	5.54	4.73	-
MMD	4.65	3.98	-

Deux modèles multitâches ont été testés: le modèle *Fully-Shared* (FS) et le modèle *Shared-Private* (SP), décrit dans ce poster. Ces deux modèles ont été comparés à des solutions monotâches (ST). Dans tous les cas de figure, les modèles multitâches ont démontré des performances accrues par rapport aux modèles monotâches. Le texte est la modalité la plus discriminante bien qu'elle puisse être combinée favorablement avec les cinq autres modalités visuelles (HP, EG, AU) et acoustiques (COV, FMT).

### CONCLUSIONS

Les résultats ne sont pas assez fiables pour aider le praticien dans son diagnostic. Pour la régression, le niveau d'erreur moyen est de l'ordre de 3.29 sur une échelle de 0 à 24. Pour la classification, le modèle garantit une précision de 60.61%. Alors que de nombreuses tâches humaines peuvent être modélisées avec succès, le domaine de la psychiatrie demeure encore un champ d'études trop peu exploré, et dont les perspectives sont immenses.

### RÉFÉRENCES

[0] Lecun, Y. 2017. L'apprentissage profond, une révolution en intelligence artificielle. Collège de France. Fayard  
 [1] Gratch, J. et al. 2014. The distress analysis interview corpus of human and computer interviews. LREC.  
 [2] Williamson, J.R. et al. 2016. Detecting depression using vocal, facial and semantic communication cues. AVEC.  
 [3] Qureshi, S.A. et al. 2019. Multitask representation learning for multimodal estimation of depression level. IEEE IS.