

تطوير نموذج هجين لكشف حالات اكتئاب اعتماداً على الإشارة الصوتية

د. م. رنيم كنج *

م. منار نجار **

(تاريخ الإيداع ٢٠٢٥/٣/٢٤. قُبل للنشر في ٢٠٢٥/٥/١٥)

□ ملخص □

يُعد الاكتئاب من الاضطرابات النفسية الأكثر انتشاراً، حيث يؤثر سلباً على جودة الحياة ويرتبط بمخاطر صحية جسيمة. وفي ظل التحديات التشخيصية التقليدية، برزت تقنيات الذكاء الاصطناعي كأدوات واعدة للكشف المبكر عن الاكتئاب عبر تحليل المؤشرات الحيوية مثل الصوت. في هذا البحث، تم تطوير نظام للكشف عن الاكتئاب السريري باستخدام الإشارات الصوتية الناطقة باللغة العربية، مع التركيز على دمج البيانات السياقية وتعزيز دقة التشخيص. تم بناء قاعدة بيانات صوتية ناطقة باللغة العربية خاصة بالاكتئاب (AADD) تضم ٧٣ متطوعاً، حيث سُجلت أصواتهم أثناء قراءة نصوص تحوي مشاعر إيجابية وسلبية وحيادية في بيئة احترافية. بالإضافة إلى التسجيلات، شملت القاعدة معلومات ديموغرافية (العمر، الجنس، مكان الإقامة، التدخين) ونتائج مقياس بيك للاكتئاب كمعيار ذهبي للتشخيص. لاستخلاص السمات الصوتية، استُخدمت خوارزمية معاملات جاماتون سيبسترال (GTCC) بالتوازي مع سمات صوتية داعمة (مثل التذبذب (Shimmer)، ومدة الصوت (Duration))، لالتقاط الأنماط المرتبطة بالحالة النفسية.

اعتمد النظام المقترح على خوارزمية (XGBoost) لتصنيف الحالات، نظراً لكفاءتها في التعامل مع البيانات متعددة الأبعاد وقدرتها على تقليل خطر التهيئة الزائدة. لضمان موثوقية النموذج، تم اختباره على مجموعتين: الأولى هي قاعدة البيانات التي تم بناؤها (AADD)، والثانية هي مجموعة البيانات المعيارية العالمية (MODMA Dataset) (الناطق باللغة الصينية). استخدم التقييم مقاييس متعددة: معدل الدقة (Accuracy)، الدقة (Precision)، معدل الاسترجاع (Recall)، القيمة (F1-Score)، دقة التحقق التبادلي (Cross-validation accuracy)، ومعدل الخطأ الكلي (Total Error Rate)، مع تقسيم البيانات بنسبة ٨٠% للتدريب و ٢٠% للاختبار في كلتا الحالتين.

أظهرت النتائج أن النظام حقق دقة بلغت ٧٧% على مجموعة (AADD)، بينما وصلت الدقة إلى ٨٥% على مجموعة (MODMA)، مما يعكس تأثير تنوع البيانات اللغوية والثقافية على الأداء. كما أبرز تحليل السمات دور التذبذب ومدة الصوت كمؤشرات بارزة للاكتئاب. تُظهر هذه الدراسة إمكانية توظيف التعلم الآلي في التشخيص النفسي، مع التأكيد على أهمية تطوير قواعد بيانات محلية لتعميم النماذج على السياقات الثقافية المتنوعة.

الكلمات المفتاحية: تحليل الإشارة الصوتية، التعلم الآلي، كشف الاكتئاب، الاكتئاب السريري، قواعد البيانات الصوتية، استخلاص السمات الصوتية، معاملات جاماتون سيبسترال (GTCC)، خوارزمية (XGBoost).

* مدرس، قسم هندسة النظم الحاسوبية والالكترونية، كلية هندسة تكنولوجيا المعلومات والاتصالات، جامعة طرطوس، طرطوس، سوريا.

raneem.knaj@gmail.com

** طالبة ماجستير، قسم هندسة النظم الحاسوبية والالكترونية، كلية هندسة تكنولوجيا المعلومات والاتصالات، جامعة طرطوس، طرطوس، سوريا.

manarnajjar7@gmail.com

Developing a hybrid model for detecting depression based on the audio signal

Dr. Eng. Raneem Knaj*

Eng. Manar Najjar**

(Received 24/3/2025 . Accepted 13/5/2025)

□ ABSTRACT □

Depression is one of the most prevalent mental health disorders, adversely affecting quality of life and associated with significant health risks. Given the limitations of traditional diagnostic methods, artificial intelligence techniques have emerged as promising tools for the early detection of depression through the analysis of biomarkers, such as voice. In this study, a system for detecting clinical depression using Arabic speech signals was developed, with a focus on integrating contextual data and improving diagnostic accuracy.

An Arabic-Audio Database for Depression (AADD) was created, comprising recordings from 73 volunteers. Their voices were recorded while reading texts containing positive, negative, and neutral emotions in a professional environment. In addition to the audio recordings, the database included demographic information (age, gender, residence, smoking status) and Beck Depression Inventory (BDI) scores as a gold standard for diagnosis. To extract audio features, the Gammatone Cepstral Coefficients (GTCC) algorithm was employed, along with supplementary audio features such as shimmer and duration, to capture patterns associated with mental states.

The proposed system utilized the XGBoost algorithm for classification, chosen for its efficiency in handling multi-dimensional data and its ability to mitigate the risk of overfitting. To ensure the model's reliability, it was evaluated on two datasets: the first one is the AADD database and the second one is the international standard dataset MODMA Dataset (in Chinese). Multiple evaluation metrics were used, including accuracy, precision, recall, F1-score, cross-validation accuracy, and total error rate, with an 80-20 split for training and testing in both cases.

The results demonstrated that the system achieved an accuracy of 77% on the AADD dataset and 85% on the MODMA dataset, highlighting the impact of linguistic and cultural diversity on performance. Feature analysis revealed that shimmer and duration were significant indicators of depression. This study underscores the potential of machine learning in psychological diagnosis and emphasizes the importance of developing localized databases to generalize models across diverse cultural contexts.

Keywords: Speech signal analysis, machine learning, depression detection, clinical depression, audio databases, feature extraction, Gammatone Cepstral Coefficients (GTCC), XGBoost algorithm.

¹ lecturer, Department of CESE, Faculty of Information and Communication Technology, Tartous University, Tartous, Syria, e-mail: raneem.knaj@gmail.com

² Master's degree student, Department of CESE, Faculty of Information and Communication Technology, Tartous University, Tartous, Syria, e-mail: manarnajjar7@gmail.com

٢- المقدمة:

يُعد الاكتئاب من أكثر الاضطرابات النفسية انتشاراً حول العالم، حيث يؤثر سلباً على جودة حياة الأفراد ويرتبط بمخاطر صحية ونفسية جسيمة، بما في ذلك زيادة خطر الإصابة بأمراض مزمنة وحتى الانتحار. وعلى الرغم من التقدم الكبير في فهم هذا الاضطراب، لا تزال عملية تشخيص الاكتئاب تواجه تحديات كبيرة، خاصةً تشخيص الاضطراب الاكتئابي الكبير أو ما يعرف بالاكتئاب السريري، لاسيما عند الاعتماد على الأساليب التقليدية مثل الاستبيانات الورقية أو المقابلات المباشرة. هذه الأساليب غالباً ما تكون عرضة لعدم الدقة، حيث قد يلجأ المريض إلى إخفاء مشاعره الحقيقية أو قد يعجز عن التعبير عنها بشكل صحيح بسبب طبيعة الاكتئاب نفسه. لذلك، هناك حاجة ماسة إلى أدوات تشخيصية أكثر موضوعية وموثوقة تعتمد على مؤشرات حيوية مثل الصوت، والتي يمكن أن توفر رؤى أعمق حول الحالة النفسية للفرد.

في هذا السياق، برزت تقنيات الذكاء الاصطناعي كأدوات واعدة في مجال التشخيص النفسي، حيث تسمح بتحليل الإشارات الصوتية واستخراج الأنماط التي تعكس الحالة النفسية للمتحدث. ومع ذلك، فإن معظم الأنظمة الحالية تعتمد على قواعد بيانات صوتية بلغات أجنبية، مما يحد من فعاليتها عند تطبيقها على الناطقين باللغة العربية، كما أنّ هنالك فقر بقواعد البيانات الناطقة باللغة العربية والخاصة بالاكتئاب. بالإضافة إلى ذلك، تعاني العديد من الأنظمة المتوفرة من دقة محدودة في الكشف عن الاكتئاب، مما يستدعي تطوير نماذج أكثر تطوراً وقدرة على التعامل مع الخصائص الفريدة للغة العربية.

هدف هذا البحث إلى معالجة هذه الفجوة من خلال بناء قاعدة بيانات صوتية ناطقة باللغة العربية خاصة بالاكتئاب. تحتوي هذه القاعدة على تسجيلات لأفراد متنوعين من حيث الحالة النفسية، مع تضمين معلومات ديموغرافية ونتائج مقياس بيك للاكتئاب كمعيار ذهبي للتشخيص. كما هدف البحث إلى تطوير نموذج متقدم لتحليل الإشارات الصوتية واستخراج السمات ذات الصلة بالاكتئاب السريري من خلال تقنيات التعلم الآلي، مما يساهم في تحسين دقة التشخيص. من خلال هذا العمل، نسعى إلى تقديم نظام تشخيصي أكثر دقة وموثوقية يمكن أن يساعد الطبيب أو المعالج النفسي في الكشف المبكر عن الاكتئاب السريري [1,2].

٣- الدراسات المرجعية ذات الصلة:

شهدت تقنيات الذكاء الاصطناعي في السنوات الأخيرة تطوراً كبيراً في مجال تحليل البيانات الصوتية، لا سيما في التشخيص الطبي والنفسي. وقد برزت هذه التقنيات كأدوات فعالة للكشف عن الاضطرابات النفسية، مثل الاكتئاب، من خلال تحليل الخصائص الصوتية وأنماط الكلام التي تعكس الحالات النفسية. في هذا الإطار، تم تطوير العديد من النماذج والخوارزميات القائمة على الذكاء الاصطناعي لتحليل الإشارات الصوتية وتصنيف حالات الاكتئاب بدقة عالية، مما عزز قدرة الأنظمة على تمييز الأنماط الصوتية المرتبطة بهذه الحالة. كما تم إنشاء قواعد بيانات صوتية متخصصة تحتوي على عينات لأفراد يعانون من اضطرابات نفسية مختلفة، مثل الاكتئاب والقلق، بالإضافة إلى عينات تعكس حالات عاطفية متنوعة كالسعادة والحزن والغضب. تُعد قواعد البيانات هذه الأساس في تدريب واختبار نماذج الذكاء الاصطناعي، حيث توفر بيانات موثوقة ومتنوعة تمكن الباحثين من تطوير أنظمة دقيقة وفعالة، مما يفتح آفاقاً جديدة في التشخيص المبكر والتدخل العلاجي وقد أجريت العديد من الدراسات البحثية في هذا المجال، نذكر أهمها:

- في دراسة Ma وآخرون، تم تقديم نموذج (DepAudioNet) وهو نموذج فعال يعتمد على التعلم العميق لتصنيف الاكتئاب بناءً على الإشارات الصوتية. تم استخدام الشبكات العصبونية الالتفافية العميقة (Deep CNN) لاستخراج الميزات الصوتية وتم تدريب النموذج على قاعدة بيانات (DAIC-WOZ)، حيث أظهر أداءً متميزاً في الكشف التلقائي عن الاكتئاب من خلال الصوت. [3]
- أما الباحث Gallardo-Antolín و Vázquez-Romero، فقد استخدموا الشبكات العصبونية الالتفافية (CNN) وتقنيات التعلم العميق لإنشاء نموذج يعتمد على دمج عدة شبكات (CNN) للكشف التلقائي عن الاكتئاب في الكلام. تم اختبار النموذج على قاعدة بيانات (DAIC-WOZ) وحقق دقة تصنيف عالية مقارنة بالنماذج التقليدية. [4]
- في بحث Smith وآخرون، تم تحليل أنماط الصوت لاكتشاف الاكتئاب بين كبار السن. اعتمد الباحثون على ميزات مشتقة من الصوت، مثل تردد النغمة والطاقة قصيرة المدى (STE)، وتم استخدام تقنيات التعلم الآلي لتحليل البيانات المسجلة من المشاركين. أظهرت النتائج أن بعض الأنماط الصوتية ترتبط بشكل واضح بحالات الاكتئاب. [5]
- بينما قدم Muzammel وآخرون نموذج (AudVowelConsNet)، وهو نموذج يعتمد على شبكة (CNN) عميقة مصممة على مستوى الفونيمات (Phoneme-level) لتشخيص الاكتئاب السريري. تم تدريب النموذج باستخدام قاعدة بيانات (DAIC-WOZ)، وحقق نتائج متميزة في تصنيف الاكتئاب اعتماداً على التحليل العميق للإشارات الصوتية. [6]
- اقترح الباحث Ye وآخرون نموذجاً متعدد الوسائط يعتمد على دمج البيانات الصوتية والنصية لتحليل الاكتئاب. استخدموا ميزات صوتية مشتقة من الكلام العاطفي جنباً إلى جنب مع تحليل النصوص التقييمية للمشاركين. تم اختبار النموذج على قاعدة بيانات (DAIC-WOZ)، وحقق أداءً جيداً في التنبؤ بمستوى الاكتئاب. [7]
- بينما اعتمد Rejaibi وآخرون على استخدام معاملات تردد الميل (MFCC) جنباً إلى جنب مع الشبكات العصبونية التكرارية (RNN) لتقييم وتشخيص الاكتئاب تلقائياً. تم تدريب النموذج على قاعدة بيانات (DAIC-WOZ)، وأظهرت النتائج أن استخدام (RNN) عزز دقة التصنيف بشكل كبير. [8]
- في بحث Sardari وآخرون، تم تقديم نموذج يعتمد على التشفير التلقائي الالتفافي (Convolutional Autoencoder) لاكتشاف الاكتئاب من خلال الإشارات الصوتية. تم استخدام قاعدة بيانات (DAIC-WOZ)، وأثبت النموذج كفاءته في استخلاص الميزات التمييزية وتحقيق نتائج تصنيف دقيقة. [9]
- أما Sun وآخرون فقد اقترحوا استخدام الشبكات العصبونية البيانية (GNN) لتحليل الإشارات الصوتية بهدف اكتشاف الاكتئاب. تم الاستفادة من التمثيلات البيانية لاستخراج العلاقات الهيكلية في الصوت، مما أدى إلى تحسين أداء التصنيف على قاعدة بيانات (DAIC-WOZ). [10]
- بينما قام Das و Naskar بتطوير نموذج يعتمد على ميزات (MFCC) وتحويل الإشارة إلى تمثيلات طيفية باستخدام الشبكات العصبونية الالتفافية (CNN) لاكتشاف الاكتئاب. تم اختبار النموذج على

مجموعات البيانات المعيارية (DAIC-WOZ) و (MODMA) و (RAVDESS)، حيث أظهر أداءً متميزاً في تصنيف حالات الاكتئاب عبر مختلف البيئات الصوتية. [11]

- أخيراً، قدم Chaudhary و Marriwala نموذجاً هجيناً يعتمد على تقنيات التعلم العميق للكشف عن الاكتئاب. تم دمج عدة نماذج شبكية لتحسين دقة التصنيف، وتم اختبار النموذج على قاعدة بيانات (DAIC-WOZ)، حيث أظهر أداءً تنافسياً مقارنة بالأساليب السابقة. [12]

٤- أهمية البحث وأهدافه:

تبع أهمية البحث من أنّ نصف حالات الانتحار المكتملة ترتبط بالاكتئاب واضطرابات المزاج الأخرى، ووفقاً لمنظمة الصحة العالمية (World Health Organization - WHO)، يقدر أن أكثر من ٣٠٠ مليون شخص يعانون من الاكتئاب، وهذا العدد آخذ في الارتفاع عالمياً [13]. ومن هنا تأتي أهمية أنظمة كشف الاكتئاب لمساعدة الأطباء والمعالجين النفسيين في تشخيص هذا الاضطراب العقلي والوقاية منه لمساعدة المرضى في أسرع وقت ممكن. لاسيما أن غالبية نظم كشف الاكتئاب المتوفرة تم تدريبها على قواعد بيانات صوتية متوفرة على مواقع الويب لناطقين بالإنكليزية مما يضعف أداءها على الناطقين باللغة العربية. يهدف هذا البحث إلى تطوير أدوات مساعدة لاكتشاف الاكتئاب السريري من خلال تحليل الإشارة الصوتية، بالاعتماد على تقنيات التعلم الآلي باستخدام خوارزمية (XGBoost). وتصميم نظام صوتي يتمتع بالميزات التالية:

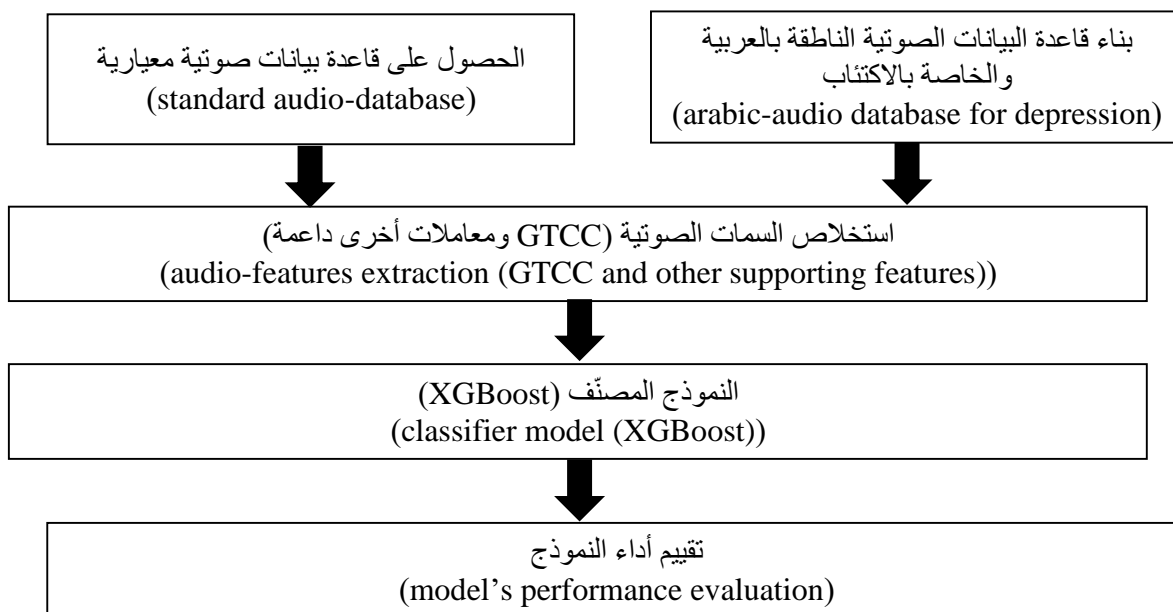
- قاعدة بيانات صوتية ناطقة باللغة العربية خاصة بالاكتئاب، تحتوي ملفات صوتية لمحدثين متنوعين، وفقاً لمقياس بيك الخاص بالاكتئاب (Beck Depression Inventory - BDI).
- وسيلة متطورة لاكتشاف الاكتئاب استناداً إلى سمات صوتية يتم استخلاصها وتمثيلها للنموذج كدخل.

٥- طرائق البحث ومواده:

يعتمد هذا البحث على المنهج التحليلي والتجريبي في عملية المقارنة بين أداء النموذج المقترح على كل من قاعدة البيانات الصوتية الناطقة باللغة العربية التي تم بناؤها، وقاعدة البيانات الصوتية المعيارية، ومناقشة النتائج. تم الاعتماد على مجموعة من الأدوات البرمجية والمادية، حيث تم الاعتماد على نظام تشغيل (windows 10)، ولغة البرمجة (python 3.11)، وبيئة البرمجة (pycharm)، كما تم استخدام كل من المكتبات (pandas)، (sklearn)، (imblearn)، (xgboost)، (pyplotlib)، (seaborn)، و (librosa) للتعامل مع الملفات الصوتية وهياكل البيانات والتعامل مع مشكلة عدم توازن العينات وبناء المصنّف (XGBClassifier) وتقييم أداء النموذج ورسم المخططات. أما بالنسبة لمجموعات البيانات فقد تم تقسيمها إلى ٨٠% تدريب و ٢٠% اختبار.

٦- النظام المقترح:

يتألف النظام المقترح من خطوات أساسية كما هو موضح في الشكل (١):



الشكل (١): وصف النظام المقترح

٦-١ الحصول على قاعدة بيانات صوتية معيارية:

تحتاج أي منهجية بحث مقترحة إلى مقارنة نتائجها مع نتائج منهجيات بحث عالمية معيارية من أجل درجة مصداقية عالية، كذلك الأمر بالنسبة لهذا البحث فالنموذج المقترح في متنته سيتم اختباره على قاعدة بيانات عالمية معيارية معتمدة وهي (MODMA Dataset) ومقارنة النتائج. حيث أن (MODMA) هي مجموعة بيانات مفتوحة المصدر متعددة الوسائط لتحليل الاضطرابات العقلية، تتضمن بيانات تخطيط كهربية الدماغ (EEG) بالإضافة لبيانات صوتية لمرضى الاكتئاب السريري. تضم البيانات الصوتية ١٥٠٨ ملف صوتي بتنسيق (wav) لـ ٥٢ متحدثاً باللغة الصينية، مسجلة بمعدل ٤٤,١ كيلو هرتز وعمق ٢٤ بت. تم تنزيل الجزء الصوتي فقط بعد التوقيع على اتفاقية ترخيص المستخدم النهائي (EULA) من الموقع الرسمي للقاعدة.

[15]

٦-٢ بناء قاعدة البيانات الصوتية الناطقة بالعربية والخاصة بالاكتئاب:

قام البحث ببناء قاعدة بيانات صوتية ناطقة باللغة العربية وخاصة بالاكتئاب، ضمت ملفات صوتية لأشخاص مختلفين من حيث الحالة النفسية، تم إطلاق الاسم (Arabic-Audio Database for Depression (AADD)) على قاعدة البيانات الصوتية المنجزة، وقد تم اعتماد المعايير العالمية في بناء قاعدة البيانات (AADD) كما يلي:

٦-٢-١ المادة اللغوية:

ضمت المادة اللغوية جمل ومجموعات كلمات تم اختيارها بعناية لضمان حملها لدلالات عاطفية واضحة ومتنوعة تغطي المشاعر الإيجابية والسلبية والمحايدة. كما تم تصميم النصوص لتتكون من كلمات شائعة الاستخدام في اللغة العربية، مما يسهل توظيفها في نظم التعرف الآلي على الكلام. قام بصياغة المادة اللغوية الأستاذ غياث حمدي كركة الحاصل على إجازة في اللغة العربية، إذ راعى في صياغتها تحقيق التوازن بين النصوص المختلفة بما يضمن تمثيلاً شاملاً لتنوع الحالات العاطفية، وهي كالتالي:

- جمل تحمل مشاعر إيجابية: (يا إلهي كم أنا سعيد! لقد حققت حلمي أخيراً)، "كم أحب لقاء أصدقائي، فهم يشعرونني بالاعتزاز والفرح!".
- جمل تحمل مشاعر سلبية: ("لا أستطيع التغلب على هذه المشاعر السلبية، أشعر بالإحباط الشديد"، "أشعر بالخوف من المستقبل، ما الذي ينتظرني؟").
- مجموعات كلمات عاطفية: إيجابية (أمان، احتواء)، سلبية (تعاسة، خيبة)، ومحايدة (تركيز، انتباه).

٦-٢-٢ الحصول على نسخة معربة من مقياس بيك:

مقياس بيك للاكتئاب (BDI) هو أداة تشخيصية شائعة تُستخدم لتقييم شدة الاكتئاب لدى الأفراد. يتكون المقياس من مجموعة من الأسئلة تغطي أعراض الاكتئاب الجسدية والنفسية، يُعتبر هذا المقياس معياراً موثقاً به في الأوساط الطبية والنفسية، حيث يساعد الأخصائيين على تحديد مستوى الاكتئاب ووضع خطة علاجية مناسبة [14]. اللغة الأصلية لمقياس بيك هي الإنجليزية، لذلك كان لابد من الحصول على نسخة معربة منه تعريباً موثقاً. تم الحصول على نسخة معربة من هذا المقياس من الدكتور أيمن حلاق المختص بالأمراض العصبية، الذي أفاد لنا بأن هذه النسخة هي المعتمدة طبياً في مشفى ابن النفيس للأمراض النفسية في دمشق كونه يعمل فيها، وبناءً عليه اعتمدنا هذه النسخة كمقياس للاكتئاب لدى المتطوعين. كما تم التواصل مع الأخصائية النفسية زين الميداني الحاصلة على إجازة في التربية قسم علم النفس والمقيمة في دمشق لتفريغ نتائج المقياس للمتطوعين وفق التالي: ٠: لا يوجد اكتئاب، ١: اكتئاب بسيط، ٢: اكتئاب متوسط، ٣: اكتئاب شديد، و ٤: اكتئاب شديد جداً.

٦-٢-٣ المتحدثون:

لضمان تمثيل متنوع ودقيق للمجموعة المستهدفة، تم اختيار ٧٣ متطوعاً (٢٧ ذكراً و ٤٦ أنثى) من طلاب جامعة طرطوس، مع مراعاة تنوع الخلفيات الاجتماعية والجغرافية وعدم وجود عيوب نطقية أو سمعية. تم ترميز المتطوعين برقم تعريفى مكون من خمس خانات يعكس الجنس، مكان الإقامة، حالة التدخين، والرقم التسلسلي. كما تم توثيق بياناتهم الديموغرافية (العمر، الجنس، مكان الإقامة، التدخين) ونتائج "مقياس بيك" في قاعدة بيانات، مما يضمن دقة التصنيف وتوفير بيانات صوتية غنية تدعم أهداف البحث.

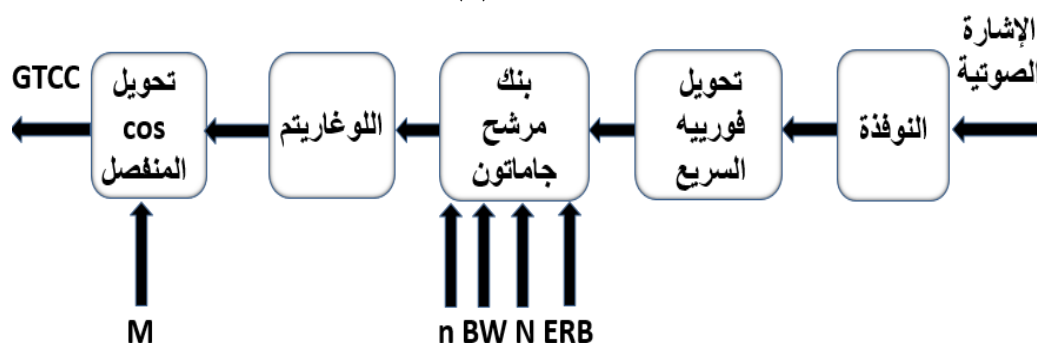
٦-٢-٤ بيئة العمل:

تم جمع البيانات الصوتية بالتعاون مع المركز الإذاعي والتلفزيوني في طرطوس، بناءً على توجيهات من جامعة طرطوس والهيئة العامة للإذاعة والتلفزيون في دمشق. تم تسجيل العينات الصوتية في استوديو معزول صوتياً ولا يوجد به تداخل كهرومغناطيسي في إذاعة طرطوس، باستخدام معدات متخصصة مثل

ميكروفون (Yamaha) وبرنامج (Adobe Audition)، وقد استغرقت التجربة بالكامل حوالي ثلاث ساعات. تم تسجيل ١٠ عينات لكل متطوع وتخزينها على حدة بتنسيق (mp3) وتسميتها من ٠ إلى ٩، ليصل المجموع إلى ٧٣٠ عينة، مع مراعاة شروط تسجيل دقيقة مثل معدل أخذ العينات ٤٤,١ كيلو هرتز وعمق ٢٤ بت. قبل التسجيل، تم توجيه المتطوعين بكيفية الجلوس والقراءة، وبعد التسجيل، طُلب منهم ملء "مقياس بيك" المعرّب، والذي يتضمن ٢١ مجموعة من العبارات لقياس حالتهم النفسية، بالإضافة إلى تقديم بياناتهم الشخصية.

٦-٣ استخلاص السمات الصوتية:

تم استخدام خوارزمية معاملات جاماتون سيسترال (Gammatone Cepstral Coefficients (GTCC)) وهي خوارزمية شائعة لاستخلاص السمات الصوتية تعتمد على نموذج معالجة القوقعة في الأذن البشرية. والتي أثبتت كفاءتها عند دمجها مع خوارزميات التعلم الآلي، إذ أظهرت تحسينات كبيرة في دقة الكشف الصوتي [16]. يتم استخلاص معاملات جاماتون سيسترال وفق الخطوات الموضحة بالشكل (٢):



الشكل (٢): خطوات حساب معاملات جاماتون سيسترال (GTCC).

يتم حساب معاملات (GTCC) وفق العلاقة التالية، حيث X_n هو طاقة الإشارة في النطاق الطيفي n ، و N هو عدد مرشحات جاماتون، و M هو عدد معاملات جاماتون سيسترال (GTCC) [17]:

$$GTCC_m = \sqrt{\frac{2}{N}} \sum_{n=1}^N \log(X_n) \cos\left[\frac{\pi n}{N} \left(m - \frac{1}{2}\right)\right] \quad 1 \leq m \leq M$$

خرج خوارزمية (GTCC) يتمثل في مجموعة من المعاملات العددية التي تمثل السمات الصوتية للإشارة المدخلة، حيث أن تطبيق تحويل جيب التمام المنفصل (DCT) على طاقات المرشحات يجعل هذه المعاملات لا ترتبط بمفاهيم فيزيائية منفصلة مثل الترددات أو الفترات الزمنية، ولكنها تمثل مكونات طيفية مضغوطة [18]. تم ضبط البارامترات كالتالي:

- عدد المرشحات ($N = 26$): اختيار هذا العدد يهدف إلى تحقيق توزيع طيفي دقيق دون فقدان التفاصيل المهمة.
- عدد المعاملات ($M = 13$): يوفر هذا العدد توازناً بين دقة التمثيل الصوتي وتقليل التحميل الحسابي.

معاملات جاماتون سيسترال تركز على تمثيل الطيف الترددي للإشارة الصوتية وتستخرج سمات تعتمد على توزيع الطاقة في النطاقات الترددية، مما يجعلها ممتازة لتمييز الأنماط الصوتية العامة. ومع ذلك، هناك سمات تكميلية متعلقة بالكتابة قد لا تغطيها (GTCC) بشكل مباشر، لذلك قمنا بحسابها أيضاً لدعم معاملات

جاماتون سيبسترال والتعزيز من شمولية السمات الصوتية في تحليل الإشارة الصوتية بالإضافة لتعزيز القدرة التنبؤية لنموذج كشف الاكثاب وتحسين دقته، وهي:

➤ التردد الأساسي (الحدة) (Pitch): هي خاصية إدراكية للصوت ترتبط بالتردد الأساسي (f_0) وتحدد ما إذا كان الصوت عالياً أو منخفضاً [19]. انخفاض أو استقرار التردد الأساسي قد يشير إلى مشاعر الحزن أو الاكثاب، حيث يقل التنوع العاطفي والنشاط الصوتي. ويتم حسابه وفق التالي:

$$R(\tau) = \sum_{n=0}^{N-\tau-1} (x(n) \cdot x(n+\tau))$$

$$f_0 = \frac{1}{T_0}$$

حيث $R(\tau)$: دالة الارتباط الذاتي عند التأخير τ ، الإشارة الصوتية، $x(n)$ التأخير أو الإزاحة الزمنية، N : طول الإشارة، T_0 : الزمن المقابل لأول قمة محلية في $R(\tau)$ ، و f_0 : هو التردد الأساسي (pitch).

➤ اضطراب التردد (Jitter): هو مقياس يعبر عن الانحراف الزمني للإشارة الصوتية عن شكلها الدوري المثالي، مما يُظهر اضطرابات في التردد الأساسي [20]. المكتتبون يظهرون زيادة في قيمة (Jitter)، مما يعكس عدم استقرار في إنتاج الصوت. يتم حسابه بالعلاقة التالية: $Jitter_{RMS}$

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (\Delta T_i - \Delta \bar{T})^2}$$

حيث: ΔT_i : الفارق الزمني بين الدورة i والدورة التالية، $\Delta \bar{T}$: متوسط الفروق الزمنية، و N : عدد الدورات.

➤ اضطراب السعة (Shimmer): يقيس التغيرات العشوائية في السعة الصوتية بين دورات الصوت المتعاقبة [21]. المكتتبون يظهرون انخفاضاً في قيم (Shimmer)، مما يعكس ضعف التحكم في السعة الصوتية، ويرتبط هذا التغير بانخفاض النشاط العاطفي والعضلي للأحبال الصوتية. يتم حسابه بالعلاقة التالية:

$$Shimmer (dB) = 20 \cdot \log_{10} \left(\frac{Amplitude_{n+1}}{Amplitude_n} \right)$$

حيث: $Amplitude_n$: السعة لدورة الصوت n ، و $Amplitude_{n+1}$: السعة لدورة الصوت التالية.

➤ مدة الصوت (Duration): هي قياس طول الزمن الذي يستغرقه الصوت من بدايته حتى نهايته [22]. المكتتبون يظهرون بطناً في الكلام مع فترات توقف أطول. يتم حسابه بالعلاقة:

$$Duration = t_{end} - t_{start}$$

حيث: t_{start} : الوقت الذي تبدأ فيه الإشارة، و t_{end} : الوقت الذي تنتهي فيه الإشارة.

➤ الطاقة الكلية (Total Energy): هي مقياس إجمالي لشدة الإشارة الصوتية يُحسب بجمع مربعات قيم السعة عبر الزمن [23]. انخفاض الطاقة الكلية يشير إلى الخمول أو قلة النشاط المرتبطة بالاكثاب، أما زيادتها قد تعكس النشاط المفرط أو التوتر. وتعطى بالعلاقة: $Energy$

$$\sum_{n=1}^N x[n]^2$$

حيث: $x[n]$: قيمة السعة في النقطة الزمنية n ، و N : إجمالي عدد النقاط الزمنية.

أخيراً، نتيجة مرحلة استخلاص السمات الصوتية هي مصفوفة أحادية البعد، تحتوي على ١٨ عنصر (سمة صوتية)، العناصر الـ ١٣ الأولى هي معاملات جاماتون سيبيسترال (GTCC)، والمعاملات الخمسة الباقية هي الطاقة الكلية (Total Energy)، الحدة (Pitch)، مدة الصوت (Duration)، اضطراب التردد (Jitter)، واضطراب السعة (Shimmer) على التوالي لكل مقطع صوتي، أي ٧٣٠ مصفوفة أحادية البعد (أشعة السمات الصوتية) لمجموعة (AADD)، و١٥٠٨ أشعة سمات صوتية لمجموعة (MODMA). أشعة السمات الصوتية هذه هي دخل المصنّف.

٦-٤ بناء النموذج المصنّف:

تم بناء نموذج تصنيف فعال لاكتشاف حالات الاكتئاب باستخدام تقنيات معالجة بيانات متقدمة، حيث تم تحسين أداء النموذج من خلال الخوارزميات التالية:

➤ بناء نموذج التصنيف باستخدام (XGBoost): تم استخدام نموذج (Extreme Gradient Boosting) لتصنيف حالات الاكتئاب بناءً على الميزات الصوتية والمعلومات الشخصية للمتطوعين. يعد (XGBoost) أحد أقوى نماذج التعلم الآلي المعتمدة على الأشجار، حيث يعتمد على (Boosting)، وهي تقنية تعتمد على تجميع عدة أشجار قرار متتالية بحيث تتعلم كل شجرة من أخطاء الشجرة السابقة، مما يعزز من دقة النموذج العام [24]. تم ضبط المعاملات الأساسية للنموذج كما يلي: عدد أشجار القرار: $(n_estimators) = 300$ ، عمق الشجرة: $(max_depth) = 5$ ، معدل التعلم: $(learning_rate) = 0.01$ ، تحديد البذرة العشوائية: $(random_state) = 42$.

➤ ترتيب السمات باستخدام المعلومات المتبادلة (Mutual Information): تم استخدام مقياس المعلومات المتبادلة (Mutual Information - MI) لحساب مدى ارتباط كل سمة من السمات المستخرجة مع الهدف (حالة الاكتئاب). حيث أن نموذج (XGBoost) حساس لترتيب أعمدة الدخل. يتم حساب المعلومات المتبادلة بين متغيرين X و Y وفقاً للعلاقة [25]:

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} \log p(x, y) \left(\frac{p(x, y)}{p(x)p(y)} \right)$$

حيث: $p(x, y)$: هو التوزيع المشترك بين المتغيرات، $p(x)$ و $p(y)$: هما التوزيعات الهامشية لكل من X و Y .

➤ موازنة العينات باستخدام تقنية (SMOTE): نظراً لعدم توازن الفئات في بيانات التدريب، حيث أن بعض مستويات الاكتئاب تحتوي على عدد أقل بكثير من العينات مقارنةً بمستويات أخرى، فقد تم استخدام تقنية (SMOTE (Synthetic Minority Over-sampling Technique لإعادة التوازن بين الفئات. تعتمد هذه التقنية على توليد عينات اصطناعية للفئة الأقل تمثيلاً من خلال الاستيفاء الخطي بين النقاط الأقرب، مما يساعد على تحسين أداء النموذج وتجنب الانحياز نحو الفئات الأكثر تمثيلاً. تم ضبط عدد الجيران ($k_neighbors$) وفقاً لأقل عدد عينات في الفئة الأصغر، بحيث يتم تجنب الأخطاء الناجمة عن اختيار عدد غير مناسب من الجيران [26]، وتم ضبطه بالشكل التالي:

$$k_neighbors = \min(5, \min \text{ samples in any class} - 1)$$

ساهمت هذه التقنيات في تحسين دقة التصنيف وتقليل معدل الخطأ، مما يعزز من كفاءة النموذج في تطبيقات الكشف عن الاكتئاب بناءً على الإشارات الصوتية.

٦-٥ تقييم أداء النموذج:

تعد عملية التقييم من المراحل الأساسية في بناء نموذج تصنيف اضطراب الاكتئاب، حيث تساعد في قياس دقة التوقعات ومدى كفاءة النموذج في التمييز بين الحالات المختلفة. تم الاعتماد على عدة معايير لتقييم أداء النموذج المستخدم، وهي:

- التحقق المتقاطع (Cross-Validation): تم استخدام تقنية (Stratified K-Fold Cross-Validation) لتقسيم البيانات إلى خمس مجموعات (طيات) متوازنة من حيث توزيع الفئات، حيث يتم تدريب النموذج على أربع طيات واختباره على الطية الخامسة، مع تكرار العملية على جميع البيانات لضمان استقرار النتائج. يوفر هذا الأسلوب تقديراً موثقاً لأداء النموذج قبل اختباره على بيانات جديدة.
- معدل الدقة (Accuracy): تعبر هذه القيمة عن نسبة التوقعات الصحيحة التي قام بها النموذج مقارنة بإجمالي عدد العينات، وهي ما يسمى بدقة التدريب. تُستخدم هذه القيمة كمؤشر عام لجودة النموذج.
- الدقة (Precision): تقيس مدى قدرة النموذج على تصنيف الحالات الإيجابية بشكل صحيح من بين جميع الحالات التي تم تصنيفها على أنها إيجابية. تعطي هذه القيمة مؤشراً على مدى موثوقية النموذج في تقديم تنبؤات دقيقة عند تصنيف حالة معينة.
- معدل الاسترجاع (Recall): يعبر عن قدرة النموذج على التعرف على جميع الحالات الإيجابية الحقيقية. يعتبر هذا المقياس مهماً في حالة التصنيف الطبي، حيث يكون من الأفضل تقليل عدد الحالات التي لم يتم اكتشافها على حساب بعض الأخطاء في التوقعات الإيجابية.
- قيمة (F1 Score): تعد هذه القيمة مقياساً متوازناً يجمع بين الدقة والاسترجاع، مما يجعلها أكثر فائدة عند التعامل مع بيانات غير متوازنة. كلما زادت هذه القيمة، كان النموذج أكثر كفاءة في تصنيف جميع الفئات.
- معدل الخطأ الكلي (Total Error Rate): يتم حساب هذه القيمة لمعرفة نسبة التوقعات الخاطئة مقارنة بإجمالي عدد العينات، مما يساعد في تقييم مدى قابلية النموذج للاعتماد في بيئات حقيقية.

٧- النتائج والمناقشة:

أظهرت النتائج أن النظام حقق معدل دقة بلغت ٧٧% على مجموعة (AADD)، بينما وصل معدل الدقة إلى ٨٥% على مجموعة (MODMA)، حيث في كلتا الحالتين تم تقسيم مجموعات البيانات إلى ٨٠% تدريب و ٢٠% اختبار.

٧-١ نتائج (AADD):

✓ تقرير التصنيف:

الجدول التالي يظهر نتائج تقرير التصنيف للنموذج على مجموعة (AADD) وفق المعايير المستخدمة:

الجدول (١): نتائج تقرير تصنيف على مجموعة (AADD).

عدد العينات (Support)	درجة (F1-score)	الاسترجاع (Recall)	الدقة (Precision)	الفئة
5	0.50	0.40	0.67	0
3	0.33	0.33	0.33	1
5	0.73	0.80	0.67	2
9	0.95	1.00	0.90	3
4	1.00	1.00	1.00	4
26	0.77	—	—	الدقة العامة
26	0.70	0.71	0.71	المتوسط الحسابي (Macro Avg)
26	0.76	0.77	0.76	المتوسط الحسابي الموزون (Weighted Avg)

تحليل الجدول (١):

• الفئة (٠): حصلت على دقة ٠,٦٧ واسترجاع ٠,٤٠، مما يعني أن النموذج يحقق تصنيفاً صحيحاً بنسبة ٦٧% عند توقع الفئة ٠، لكنه لم ينجح في استرجاع جميع العينات الفعلية من هذه الفئة (٤٠%).

• الفئة (١): أدائها ضعيف، حيث كانت الدقة والاسترجاع ٠,٣٣ فقط، مما يشير إلى أن النموذج غير قادر على التمييز الجيد لهذه الفئة.

• الفئة (٢): تحقق أداءً جيداً نسبياً، حيث بلغ الاسترجاع ٠,٨٠، مما يعني أن ٨٠% من العينات الفعلية صُنِّفت بشكل صحيح، بينما الدقة كانت ٠,٦٧.

• الفئة (٣): حققت أداءً ممتازاً مع دقة ٠,٩٠ واسترجاع ١,٠٠، مما يدل على أن جميع العينات الفعلية للفئة ٣ تم التعرف عليها بشكل صحيح، مع وجود بعض الأخطاء الطفيفة في تصنيف العينات الأخرى ضمن هذه الفئة.

• الفئة (٤): تحقق أداءً مثالياً، حيث كانت جميع القيم ١,٠٠، مما يعني أن جميع العينات الخاصة بهذه الفئة قد تم تصنيفها بشكل صحيح دون أي أخطاء.

• الدقة العامة (Accuracy) = 0.77: تشير إلى أن ٧٧% من العينات تم تصنيفها بشكل صحيح من إجمالي العينات البالغ عددها ٢٦.

• المتوسط الحسابي (Macro Avg): يتم حساب هذا المتوسط من خلال أخذ متوسط بسيط (غير موزون) لجميع القيم عبر الفئات المختلفة. في هذه الحالة، فإن: الدقة الكلية (Precision) = 0.71، الاسترجاع الكلي (Recall) = 0.71، قيمة F1 الكلية = ٠,٧٠.

• المتوسط الحسابي الموزون (Weighted Avg): يتم حسابه بأخذ المتوسط المرجح لكل مقياس (الدقة، الاسترجاع، وقيمة F1) بناءً على عدد العينات في كل فئة. القيم هي: الدقة الموزونة (Precision) = 0.76، الاسترجاع الموزون (Recall) = 0.77، قيمة F1 الموزونة = ٠,٧٦. هذا المتوسط أكثر

تعبيراً عن الأداء الفعلي للنموذج، حيث يأخذ في الاعتبار حجم الفئات، مما يعني أن الفئات ذات العينات الكبيرة لها تأثير أكبر على القيمة الإجمالية.

✓ التحقق المتقاطع:

بلغت قيمة دقة التحقق المتقاطع (Cross-validation accuracy) = 0.6243 أي أن متوسط الدقة التي حققها النموذج على عدة تقسيمات مختلفة من البيانات هو ٦٢,٤٣%. هذه النتيجة تعني أن النموذج يؤدي بشكل مقبول، مما يعني أن النموذج قابل للتعميم بشكل متوسط.

✓ معدل الخطأ الكلي:

بلغ معدل الخطأ الكلي (Total Error Rate) = 0.1154، أي أن معدل الخطأ ١١,٥٤%، وهي قيمة جيدة وتعني أن النموذج نجح في تصنيف ٨٨,٤٦% من الحالات بشكل صحيح.

٧-٢ نتائج (MODMA):

✓ تقرير التصنيف:

الجدول التالي يظهر نتائج تقرير التصنيف للنموذج على مجموعة (MODMA) وفق المعايير المستخدمة:

الجدول (٢): نتائج تقرير تصنيف على مجموعة (MODMA).

عدد العينات (Support)	F1-درجة (F1 score)	الاسترجاع (Recall)	الدقة (Precision)	الفئة
6	0.67	0.67	0.67	0
6	0.92	1.00	0.86	1
6	0.91	0.83	1.00	2
5	0.80	0.80	0.80	3
3	1.00	1.00	1.00	4
26	—	—	0.85	الدقة العامة
26	0.86	0.86	0.86	المتوسط الحسابي (Macro Avg)
26	0.85	0.85	0.85	المتوسط الحسابي الموزون (Weighted Avg)

تحليل الجدول (٢):

- الفئة (٠): حصلت على دقة ٠,٦٧ واسترجاع ٠,٦٧، مما يعني أن النموذج يحقق تصنيفاً صحيحاً بنسبة ٦٧% عند توقع الفئة ٠، وهو تحسن عن النتيجة السابقة.
- الفئة (١): أداء ممتاز، حيث بلغت الدقة ٠,٨٦ والاسترجاع ١,٠٠، مما يعني أن جميع العينات الفعلية من الفئة ١ تم استرجاعها، لكن هناك نسبة خطأ طفيفة في الدقة.

• الفئة (٢): تحقق أداءً جيداً جداً، حيث بلغ الاسترجاع ٠,٨٣، مما يعني أن ٨٣% من العينات الفعلية صُنِّفت بشكل صحيح، بينما الدقة كانت ١,٠٠، مما يشير إلى عدم وجود أي عينات خاطئة عند تصنيف هذه الفئة.

• الفئة (٣): أداء جيد، حيث حققت دقة ٠,٨٠ واسترجاع ٠,٨٠، مما يدل على توازن بين القدرة على استرجاع العينات وتصنيفها بشكل دقيق.

• الفئة (٤): تحقق أداءً مثالياً، حيث كانت جميع القيم ١,٠٠، مما يعني أن جميع العينات الخاصة بهذه الفئة تم تصنيفها بشكل صحيح دون أي أخطاء.

• الدقة العامة (Accuracy) = 0.85: تعني أن ٨٥% من العينات تم تصنيفها بشكل صحيح من إجمالي العينات البالغ عددها ٢٦.

• المتوسط الحسابي (Macro Avg): يأخذ متوسطاً بسيطاً لجميع القيم عبر الفئات المختلفة. القيم هي: الدقة الكلية (Precision) = 0.86، الاسترجاع الكلي (Recall) = 0.86، قيمة F1 الكلية = ٠,٨٦. هذا يشير إلى توازن عام جيد في أداء النموذج عبر جميع الفئات.

• المتوسط الحسابي الموزون (Weighted Avg): قيمه كانت كالتالي: الدقة الموزونة (Precision) = 0.85، الاسترجاع الموزون (Recall) = 0.85، درجة F1 الموزونة = ٠,٨٥. هذا يشير إلى أن النموذج يعمل بشكل مستقر مع تأثير حجم الفئات المختلفة.

✓ التحقق المتقاطع:

بلغت قيمة دقة التحقق المتقاطع (Cross-validation accuracy) = 0.7686 أي أن متوسط الدقة التي حققها النموذج على عدة تقسيمات مختلفة من البيانات هو ٧٦,٨٦%. هذه النتيجة تعني أن النموذج يؤدي بشكل جيد جداً، مما يعني أن النموذج قابل للتعميم بشكل جيد جداً.

✓ معدل الخطأ الكلي:

بلغ معدل الخطأ الكلي (Total Error Rate) = 0.385، أي أن معدل الخطأ ٣,٨٥%، وهي قيمة ممتازة وتعني أن النموذج نجح في تصنيف ٩٦,١٥% من الحالات بشكل صحيح.

٣-٧ مقارنة نتائج (AADD) و (MODMA):

الجدول التالي يظهر مقارنة بين نتائج كل من المجموعتين (AADD) و (MODMA) على النموذج المقترح وفق معايير التقييم المستخدمة:

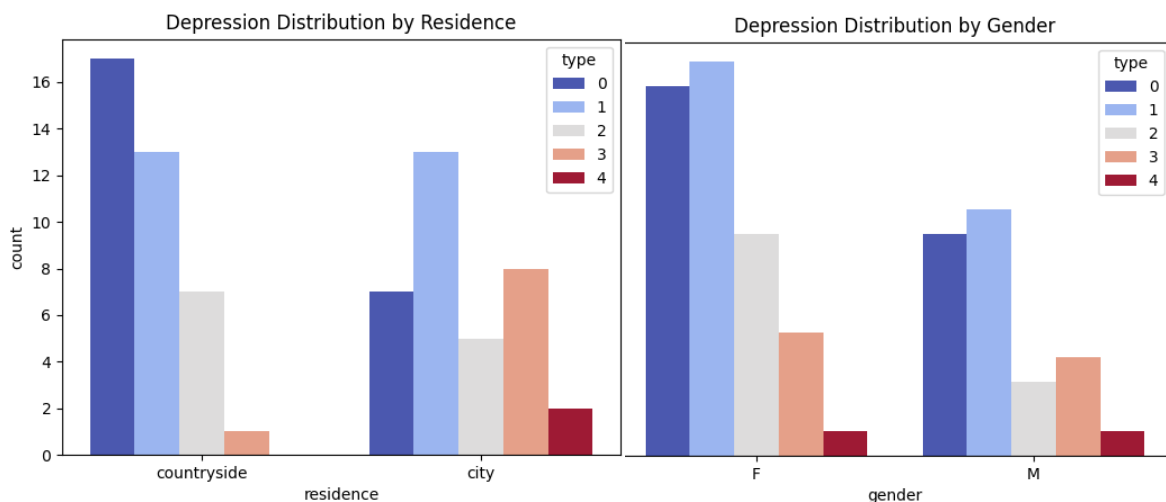
الجدول (3): مقارنة نتائج (AADD) و (MODMA).

معدل الخطأ الكلي (total error rate)	دقة التحقق المتقاطع (cross-validation accuracy)	قيمة F1 (F1 score)	معدل الاسترجاع (recall)	الدقة (precision)	معدل الدقة (accuracy)	معياري التقييم المجموع
%١١,٥٤	%٦٢,٤٣	%٧٦	%٧٧	%٧٦	%٧٧	(AADD)
%٣,٨٥	%٧٦,٨٦	%٨٥	%٨٥	%٨٥	%٨٥	(MODMA)

النتائج السابقة توضح أن النموذج أظهر نتائج جيدة جداً على مجموعة البيانات المعيارية (MODMA)، مما يضمن موثوقيته.

٧-٤ تحليل البيانات الديموغرافية:

من أجل تحليل البيانات الديموغرافية في مجموعة (AADD)، لدينا المخططات التالية التي تظهر تحليل الاكتئاب وفق الجنس ومكان الإقامة وحالة التدخين والعمر:



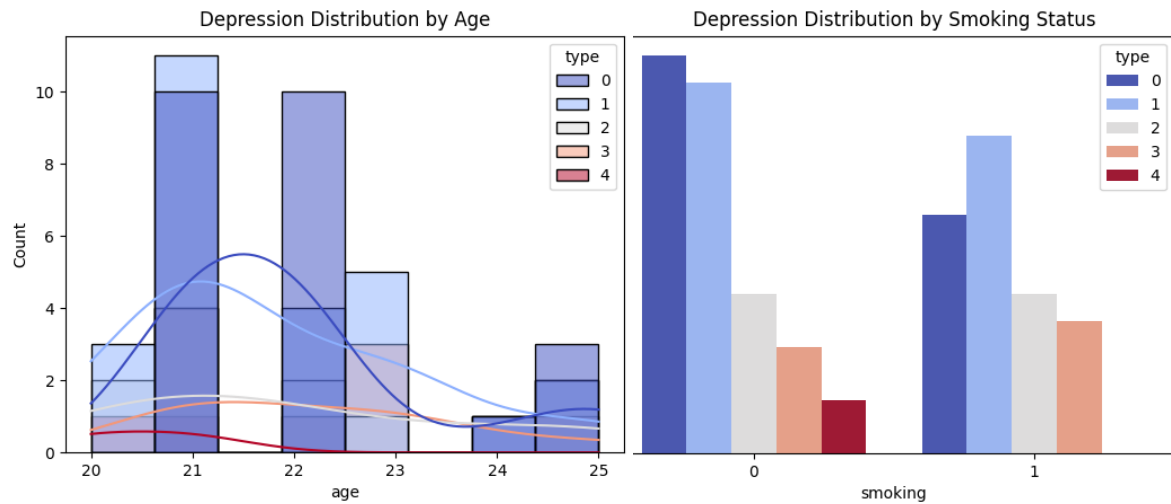
الشكل (٤): تحليل الاكتئاب وفق مكان الإقامة.

الشكل (٣): تحليل الاكتئاب وفق الجنس.

يظهر الشكل (٣) أن نسبة الإصابة بالاكتئاب تختلف حسب الجنس. يلاحظ أن كلا الجنسين يعانيان من الاكتئاب بمستويات متفاوتة، لكن يبدو أن هناك نسبة أعلى من الاكتئاب في إحدى الفئات عند الإناث مقارنة بالذكور، ويمكن تفسير هذه النتيجة بعوامل اجتماعية أو بيولوجية تؤثر على القابلية للإصابة بالاكتئاب السريري من الفئة الأولى والثانية (البسيط والمتوسط) لدى الإناث أكثر من الذكور.

يظهر الشكل (٤) أن الأفراد الذين يعيشون في المدن والريف لديهم مستويات اكتئاب مختلفة. هناك اختلاف في توزيع

الفئات، حيث يمكن ملاحظة أن فئات الاكتئاب الثالثة والرابعة (الشديد والشديد جداً) أكثر شيوعاً في المدن مقارنة بالريف والعكس صحيح للفئة الثانية (الاكتئاب المتوسط). قد يكون نمط الحياة، الضغوط المعيشية، والتفاعل الاجتماعي دور في هذه الاختلافات.



الشكل (٦): تحليل الاكتئاب وفق

الشكل (٥): تحليل الاكتئاب وفق حالة التدخين.

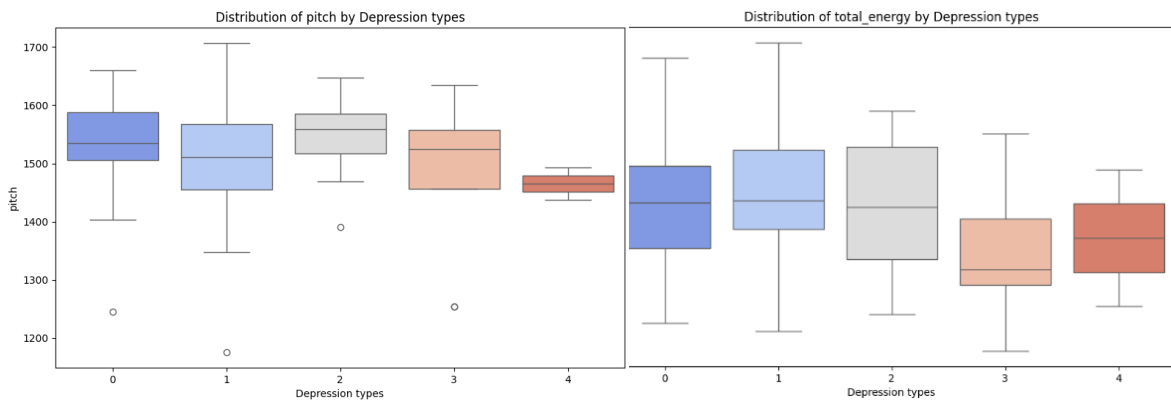
العمر.

يظهر الشكل (٥) أن الأفراد الغير مدخنين (فئة ٠) أقل عرضة للإصابة بالاكتئاب عامةً (الفئة ٠)، بينما الأفراد المدخنين أكثر عرضة للإصابة بالاكتئاب الشديد (فئة ٣)، مما يعكس التأثير السلبي للتدخين على الصحة النفسية.

يظهر الشكل (٦) أن توزيع حالات الاكتئاب بين الأعمار من ٢٠ إلى ٢٥ يعكس تبايناً ملحوظاً، إذ يمكن ملاحظة أن الأفراد في سن ٢١ و ٢٢ يمثلون أعلى معدلات الإصابة بالاكتئاب، حيث تتركز الأعداد الكبيرة للفئات المختلفة للاكتئاب. في المقابل، تنخفض معدلات الاكتئاب بشكل واضح في الأعمار ٢٤ إلى ٢٥. تشير هذه البيانات إلى أن سن العشرينات المبكرة قد يكون فترة تتأثر فيها القابلية للإصابة بالاضطرابات النفسية، مما قد يرتبط بالضغط النفسي المتزايد للأفراد خلال هذه الفترة.

٥-٧ تحليل السمات الصوتية الداعمة:

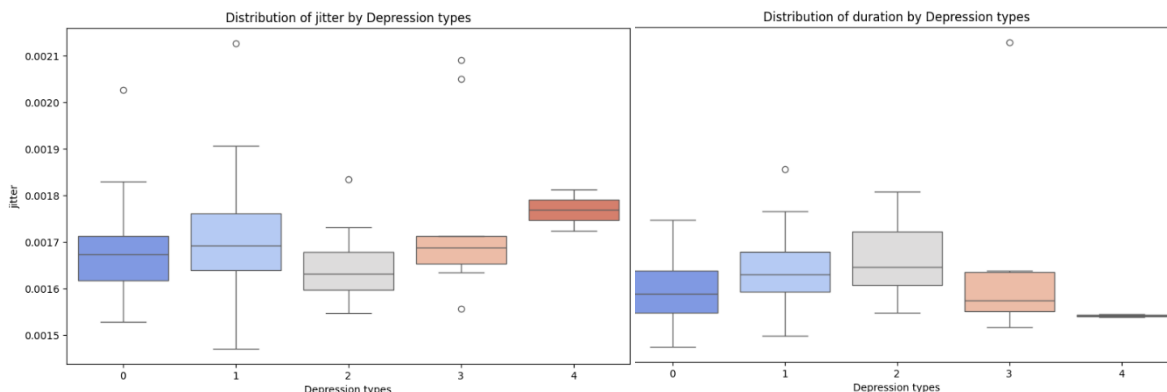
من أجل تحليل السمات الصوتية الداعمة في مجموعة (AADD)، تظهر المخططات التالية توزيع الاكتئاب وفق الطاقة الكلية (Total Energy)، الحدة (Pitch)، مدة الصوت (Duration)، اضطراب التردد (Jitter)، واضطراب السعة (Shimmer):



الشكل (٨): توزيع الاكتئاب وفق

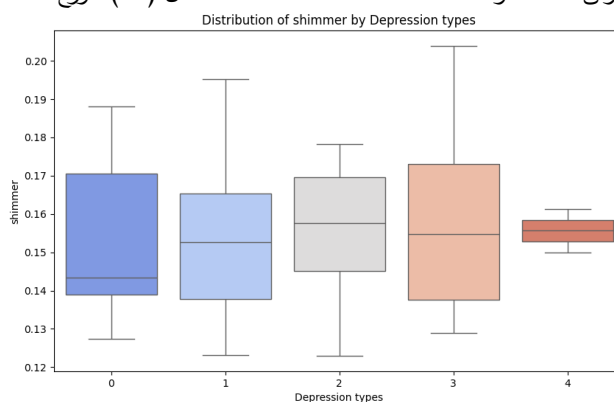
الشكل (٧): توزيع الاكتئاب وفق الطاقة الكلية.

الحدة.



الشكل (١٠): توزيع الاكتئاب وفق اضطراب التردد.

الشكل (٩): توزيع الاكتئاب وفق مدة الصوت.



الشكل (١١): توزيع الاكتئاب وفق اضطراب السعة.

يظهر الشكل (٧) أن الطاقة الكلية تقل بشكل ملحوظ مع زيادة شدة الاكتئاب. الأفراد الذين يعانون من الاكتئاب الشديد جداً (الفئة ٤) تمتلك أقل مؤشرات الطاقة، مما قد يرتبط بانخفاض الحافز والنشاط في الحياة اليومية. الفئات الأخرى تحتفظ بمستويات طاقة أعلى، مما يعكس القدرة على الاستجابة والتفاعل الاجتماعي بشكل أفضل.

يوضح الشكل (٨) اختلافات واضحة في الحدة بين الفئات المختلفة. الأفراد المصابون بدرجات الاكتئاب العالية (الفئات ٣ و ٤)، لديهم حدة أقل، مما يشير إلى أن الاكتئاب الشديد والمستويات العاطفية السلبية تؤثر على نغمة الصوت. بينما يميل الأفراد الذين لا يعانون من الاكتئاب إلى حدة أعلى، مما يدل على توازن عاطفي أكبر.

يظهر الشكل (٩) أن مدة الصوت تختلف باختلاف أنواع الاكتئاب. يتضح أن الأفراد الذين يعانون من الاكتئاب الشديد جداً (الفئة ٤) لديهم مدة صوت أقل مقارنة بالفئات الأخرى، مما يعني أنهم يميلون للتحدث بسرعة عالية. بينما تظهر الفئات الأخرى (٠-٣) تبايناً أقل في تقلبات المدة، مما قد يدل على أن التأثيرات السلبية للاكتئاب على الصوت تتزايد مع شدة الحالة.

ينبئ الشكل (١٠) عن وجود ارتفاع طفيف في اضطراب التردد لدى الفئات الأكثر اكتئاباً. يمكن ملاحظة أن الأفراد الذين يعانون من الاكتئاب الشديد والشديد جداً (الفئات ٣ و ٤) لديهم قيم أعلى من اضطراب

التردد مقارنة بالفئات الأخرى. يُعزى هذا إلى زيادة القلق وما يترتب على الاكتئاب من توتر نفسي يؤثر على استقرار الصوت.

يتضح من الشكل (١١) أن هناك تبايناً واضحاً في قيمة اضطراب السعة بين أنماط الاكتئاب، إذ أن الأفراد الذين يعانون من اكتئاب شديد جداً (الفئة ٤) يظهرون مستويات أقل من اضطراب السعة، مما قد يعني فقدان القوة والتعبير في الصوت كنتيجة لأعراض الاكتئاب. بينما الفئات الأخرى تبدي قيم أكبر من اضطراب السعة، مما يدل على مرونة أكبر في الصوت.

٨- التوصيات:

هناك مجموعة من التوصيات المستقبلية تشير إلى الإجراءات والتحسينات التي يمكن اتخاذها لتطوير النموذج وتحسين أدائه في كشف حالات الاكتئاب. منها:

١- تحسين دقة النموذج: يمكن تجربة خوارزميات ذكاء صناعي إضافية مثل الشبكات العصبية العميقة أو النماذج الهجينة التي تجمع بين الشبكات العصبونية والتعلم التقليدي لتعزيز الدقة. كما يمكن تحسين ضبط (Hyperparameters) لنموذج (XGBoost) باستخدام تقنيات مثل (Grid Search) أو (Bayesian Optimization).

٢- معالجة عدم توازن البيانات: بالرغم من استخدام تقنية (SMOTE) لموازنة الفئات، يمكن تجربة تقنيات أخرى مثل (ADASYN) أو (Class Weighting) لاختبار تأثيرها على دقة النموذج. كما يمكن استخدام مجموعات بيانات أكبر تحتوي على تمثيل أكثر توازناً للفئات المختلفة.

٣- تحليل السمات الصوتية بعمق: يمكن استكشاف ميزات إضافية من الإشارة الصوتية، مثل (Mel-Frequency Cepstral Coefficients (MFCCs) أو (Spectral Contrast)، التي قد تعزز دقة النموذج. كما يمكن دراسة تأثير معدل العينة (Sampling Rate) على دقة استخراج الميزات الصوتية.

٤- تحليل العوامل الديموغرافية وتأثيرها على الاكتئاب: أظهرت النتائج وجود فروق في انتشار الاكتئاب بين الجنسين ومناطق السكن، لذا يمكن دمج معلومات ديموغرافية أخرى مثل المستوى التعليمي أو الوضع الاقتصادي للحصول على نموذج أكثر تفسيراً. كما يمكن تحليل العلاقة بين التدخين والاكتئاب بطرق إحصائية متقدمة، مثل تحليل التباين (ANOVA).

٥- تطبيق عملي للنموذج: يمكن نشر النموذج في شكل تطبيق ويب أو تطبيق هاتف ذكي يوفر أداة مساعدة للأخصائيين النفسيين لتقييم الاكتئاب باستخدام التسجيلات الصوتية. ويمكن أيضاً التعاون مع المؤسسات الصحية لتوظيف النموذج كأداة دعم في التشخيص المبكر للاكتئاب، مما قد يساعد في التدخل المبكر والعلاج الفعال.

٦- توسيع نطاق الدراسة: يمكن توسيع نطاق البيانات لتشمل لغات ولهجات أخرى للتأكد من تعميم النموذج على نطاق أوسع.

٩ - المراجع المستخدمة:

- [1] American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders, 5th edn edn. American Psy-chiatric Association., Retrieved from.
- [2] Taschereau-Dumouchel, V., Cortese, A., & Kawato, M. (2022). Artificial intelligence for mental health and mental illnesses: an overview. *Current Opinion in Behavioral Sciences*, 40, 69-84. <https://doi.org/10.1016/j.cobeha.2021.12.007>
- [3] Ma, X., Yang, H., Chen, Q., Huang, D., & Wang, Y. (2016, October). Depaudionet: An efficient deep model for audio based depression classification. In Proceedings of the 6th international workshop on audio/visual emotion challenge (pp. 35-42).
- [4] Vázquez-Romero, A., & Gallardo-Antolín, A. (2020). Automatic detection of depression in speech using ensemble convolutional neural networks. *Entropy*, 22(6), 688.
- [5] Smith, M., Dietrich, B. J., Bai, E. W., & Bockholt, H. J. (2020). Vocal pattern detection of depression among older adults. *International journal of mental health nursing*, 29(3), 440-449.
- [6] Muzammel, M., Salam, H., Hoffmann, Y., Chetouani, M., & Othmani, A. (2020). AudVowelConsNet: A phoneme-level based deep CNN architecture for clinical depression diagnosis. *Machine Learning with Applications*, 2, 100005.
- [7] Ye, J., Yu, Y., Wang, Q., Li, W., Liang, H., Zheng, Y., & Fu, G. (2021). Multi-modal depression detection based on emotional audio and evaluation text. *Journal of Affective Disorders*, 295, 904-913.
- [8] Rejaibi, E., Komaty, A., Meriaudeau, F., Agrebi, S., & Othmani, A. (2022). MFCC-based recurrent neural network for automatic clinical depression recognition and assessment from speech. *Biomedical Signal Processing and Control*, 71, 103107.
- [9] Sardari, S., Nakisa, B., Rastgoo, M. N., & Eklund, P. (2022). Audio based depression detection using Convolutional Autoencoder. *Expert Systems with Applications*, 189, 116076.
- [10] Sun, C., Jiang, M., Gao, L., Xin, Y., & Dong, Y. (2024). A novel study for depression detecting using audio signals based on graph neural network. *Biomedical Signal Processing and Control*, 88, 105675.
- [11] Das, A. K., & Naskar, R. (2024). A deep learning model for depression detection based on MFCC and CNN generated spectrogram features. *Biomedical Signal Processing and Control*, 90, 105898.
- [12] Marriwala, N., & Chaudhary, D. (2023). A hybrid model for depression detection using deep learning. *Measurement: Sensors*, 25, 100587.
- [13] World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates; Technical Report; World Health Organization: Geneva, Switzerland, 2017.
- [14] Beck, A.T.; Steer, R.A.; Carbin, M.G. Psychometric properties of the Beck Depression Inventory: Twenty-fiveyears of evaluation. *Clin. Psychol. Rev.* 1988, 8, 77–100.
- [15] Cai, H., Gao, Y., Sun, S., Li, N., Tian, F., Xiao, H., Li, J., Yang, Z., Li, X., Zhao, Q., Liu, Z., Yao, Z., Yang, M., Peng, H., Zhu, J., Zhang, X., Hu, X., & Hu, B. (2020). MODMA dataset: a Multi-modal Open Dataset for Mental-disorder Analysis. arXiv preprint arXiv:2002.09283
- [16] Liu, G. K. (2018). Evaluating Gammatone Frequency Cepstral Coefficients with Neural Networks for Emotion Recognition from Speech. arXiv preprint arXiv:1806.09010.
- [17] Valero, X., & Alías, F. (2012). Gammatone Cepstral Coefficients: Biologically Inspired Features for Non-Speech Audio Classification. *IEEE Transactions on Multimedia*, 14(6), 1684-1689.
- [18] Boualoulou, N., Drissi, T. B., & Nsiri, B. (2023). CNN and LSTM for the classification of parkinson's disease based on the GTCC and MFCC. *Applied Computer Science*, 19(2), 1-24

- [19] Chen, Y., Jiang, H., Li, C., Jia, X., & Ghamisi, P. (2016). Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10), 6232-6251.
- [20] Kuncheva, L. I., & Faithfull, W. J. (2014). PCA feature extraction for change detection in multidimensional unlabeled data. *IEEE transactions on neural networks and learning systems*, 25(1), 69- 80.
- [21] Titze, I. R. (1994). *Principles of Voice Production*. Prentice Hall.
- [22] Ghamisi, P., Benediktsson, J. A., & Sveinsson, J. R. (2014). Automatic Spectral-Spatial Classification Framework Based on Attribute Profiles and Supervised Feature Extraction. *IEEE Trans. Geoscience and Remote Sensing*, 52(9), 5771-5782.
- [23] Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, 71, 10-49.
- [24] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [25] Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Wiley. <https://doi.org/10.1002/047174882X>
- [26] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357. <https://doi.org/10.1613/jair.953>