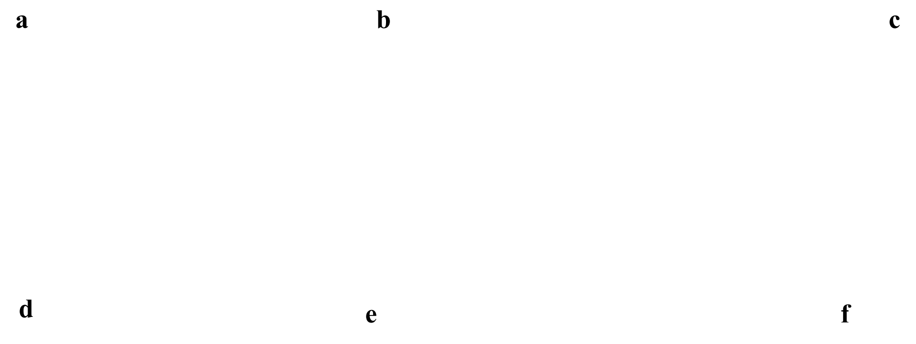
Briefings in Bioinformatics, 2025, **26(1)**, bbae554
https://doi.org/10.1093/bib/bbae554
Problem Solving Protocol

Toward molecular diagnosis of major depressive disorder by plasma peptides using a deep learning approach

Jiaqi Wang[1,2,†], Ronggang Xi[3,†], Yi Wang[2,4], Huiyuan Gao[1,*], Ming Gao

---

Received: April 22, 2024. **Revised:** September 30, 2024. **Accepted:** October 1, 2024



F 1. S MDD LC MS , ( ) MS DL. F , ( ) MDD HC , ( ) , F , ( ) CMS-N , ( ) ( )

, MS ML DL [16]. I , DL . H LC MS [17, 18]. I MS (CNN) [19]. H . I DL MS ML [20]. T -CNN DL MS MDD I CMS-N MDD. . A LC MS CMS-N ML K- (KNN), (RF), (SVM), (LDA), G B (GB) DL (LSTM, R N , T). I , LC MS . T CMS-N MDD LC MS, CMS-N (F . 1). T . T [21] P ' [22] , .

Materials and methods

Data sources

T (.) [12]. A 118 MDD F32 F33 ICD-10 ICD-10 MDD. A 120 - D C H - (HC). A MDD 18 75 HC . T E C D S P H D H . T 238 A ($n = 120$, / = 60/60) B ($n = 118$, / = 43/75). T A 60 HC (46.6 ± 15.4) 60 MDD (51.9 ± 14.2 =36/24, / =26/34). T B 60 HC (45.5 ± 9.9) 58 MDD (51.9 ± 14.4 , =39/19, / =29/29) (T S1). B 7 . . 8 . . T A MDD . A B LC MS. T . [12]. B , . F >50% . T

Batch removal and cluster analysis

T ACC, . T . T R-4.3.0 RS -2023.03 . U C B SVA RS 23, 24 . F , H (:// . . /) -SNE . T R (PCA) . T SVA . C PCA , - .

Deep learning methods

T 3.8, K 2.4.3, S - 1.3.2, T F 2.4.0, GPU. CMS-N SN 21 LC, MS . T DL 1.2 S .

T - 1D (. . ; _ 1D).

$$X\ I = X_I^1, X_I^2, \ldots.. X_I^n \tag{1}$$

$$X\ \ lc1D = LocalConnect1D\,(X\ \ I)\,, \tag{2}$$

I I, {0,1}. T : =16, - =32, =3. T : =32, - =64, =5. T - MLP (X MLP) (X MHAttention). T - MLP :

$$X\ \ MLP = Dropout\,(relu\,(W_r X_l)) + Dropout\,(\tanh\,(W_t X_l))\,, \tag{3}$$

X_l , W_r, W_t .

$$\mathrm{X}\ \ MHAttention = MultiHeadAttention\,(X\ \ MLP) \tag{4}$$

D , , , .

$$Query = query\ \ dense\,(X\ \ MLP) \tag{5}$$

$$Key = key\ \ dense\,(X\ \ MLP) \tag{6}$$

$$Value = value\ \ dense\,(X\ \ MLP)\,, \tag{7}$$

$Query$, Key $Value$ - . T :

$$Linear(X) = X \cdot W + B, \tag{8}$$

X, , B . T :

$$Attention\,(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \times V, \tag{9}$$

Q , K , V , d_k . T :

$$Concatenated\ \ Output = Concatenate\,(Attention_1, \ldots, Attention_m) \tag{10}$$

$$MultiHeadAttention\ \ Output = Output\ \ dense\,(Concatenated\ \ Output)\,, \tag{11}$$

m , 4, $Output\ \ dense$.

Machine learning methods

A ML 3.7. RF, LDA, SVM, KNN GB S - 1.0.2 25 . T DL . G S CV 26 - - 27 , .

Evaluation metric diagnosis

T DL ML , ACC (AUC). T HC , , (ROC) . ACC , . T ROC AUC (ROC-AUC) - . S ROC (TPR) (FPR), 25, 28 . T ACC AUC ACC AUC 10 - . S - HC , - . I . F , t- 29

CMS-N ML DL *P*- ≤.05.

Feature weight analysis

T 30 (), T T . P PEAKS 7.0 (BSI, C) MS/MS T P S 7.0 (BSI, C) . T H U (20 422) . T 10 0.05 D . N V PTM D B E, Q, (N-) . T PTM 5. E . T FDR PEAK <1%. T 100 PEAKS (10 , RT 2).

Compare with other DL methods

LSTM [31, 32], R N [33] T [34] MDD . T DL CMS-N, CMS-N LSTM, R N T . A

Results

Batch removal effect

T MDD LC, MS MS MDD . I A B MDD -SNE . A (F . 2 F . S1), A B . T PCA MDD (F . S1). T SVA . I (F . S1). T SVA (F . 2), PCA . A A

B .

Comparison with CMS-Net classification performance evaluation

CMS-N (F . 3) MDD . I . T CMS-N (F . 3). F MDD 30%, . F . T (4:1). F (; F . S2). T CMS-N

U CMS-N . O ACC 0.9441 AUC 0.9634. A (F . 3), 0.9352 0.9517 . T ACC AUC 0.9630 0.9931 (F . 4). T

N DL , LSTM, T R N , 10 LSTM MDD ACC 0.6533 AUC 0.7782. S ML RF, LDA, SVM, KNN, GB (10). A *P*- *t*- CMS-N (T 1). I ACC AUC (F . 4, F . S3). B CMS-N

CMS-Net weight analysis screening major depressive disorder important features

G A ( ), 3D,

a

b

c

F 2. B MDD A, . ( ) -SNE 2 3 MDD A B MDD SVA . G 0 1 B ( ) PCA ( )

. C SH A P ( ) 35, 37, 38, 39, 40, 41 . T 2D, H . U F I 28 14 PEAKS (F . 5). T (T 2), N- 2D, -1- -2-HS- -1 C3, C4-B, A-I, (T S2). C (C3, C4-B), -2-HS- -1 . F A-I, . T

(α, β, δ) A A N- 2D, T . S

Discussion

CMS-N . T LC, MS HC . C ACC . M LC, MS P MDD a

b

c

F 3. CMS-N (, 1D , ,), . ( ) T (- MLP), ( ), . ( ) C M 1: CMS-N ; 2: ; 3: . ( ) C .

T 1. C CMS-

Dataset	Classes	Model	ACC	AUC	Sensitivity	Specificity	*P*-value
MDD	2	CMS-N	0.9441 ± 0.02	0.9634 ± 0.01	0.9352 ± 0.03	0.9517 ± 0.05	
		LSTM	0.6533 ± 0.10	0.7782 ± 0.06	0.7248 ± 0.13	0.5917 ± 0.25	7.59311E-14
		T	0.8330 ± 0.09	0.8951 ± 0.08	0.8464 ± 0.07	0.8214 ± 0.14	7.97501E-11
		R N	0.9070 ± 0.04	0.9605 ± 0.02	0.8960 ± 0.06	0.9166 ± 0.05	6.62014E-05
		RF	0.8589 ± 0.04	0.9537 ± 0.02	0.8336 ± 0.06	0.8807 ± 0.06	4.01758E-13
		LDA	0.8296 ± 0.07	0.8855 ± 0.04	0.768 ± 0.09	0.8828 ± 0.07	5.74276E-27
		SVM	0.8519 ± 0.03	0.9004 ± 0.02	0.8000 ± 0.03	0.8966 ± 0.05	2.82166E-25
		KNN	0.8222 ± 0.03	0.9364 ± 0.02	0.9200 ± 0.03	0.73798 ± 0.06	1.86185E-27
		GB	0.8889 ± 0.02	0.9623 ± 0.01	0.8720 ± 0.02	0.9034 ± 0.03	2.81307E-21

Evaluation metrics: ACC, AUC values, sensitivity, and specificity (mean ± standard deviation of ten independent iterations).

T 2. T MDD

Accession	Peptide	−10lgP	Feature total contribution degree
P69905 HBA_HUMAN	VLSPADKTNVKAA GKVGAHAGE GAEALERMF	63.51	7.77733E-07
P68871 HBB_HUMAN	VHLTPEEKSAVT(−18.01)	32.67	7.65409E-07
P68871 HBB_HUMAN	VHLTPEEKSAVTAL	81.71	4.81579E-07
P69905 HBA_HUMAN	VLSPADKTNVKAA GKVGAHAGE(+37.96)	41.57	4.44261E-07
P02042 HBD_HUMAN	LVV P TQRF	53.26	2.91214E-07
P69905 HBA_HUMAN	VLSPADKTNVKAA GKVGAHAGE GAEAL	76.17	2.82464E-07
P69905 HBA_HUMAN	LASVSTVLTSK R	60.82	2.40603E-07

F 4. ROC - - .

H

ACC MDD . F , K *et al.*

- MDD -MDD AUC 0.892, 83.9%, 93% [42]. O RNA R-542-3 , R-181 - MDD 3 , R-3690, MDD AUC 0.67 [43]. D ACC MDD [44].

T . T , [45, 46]. C LC MS MDD , . ' [47] [48, 49]. T [50, 51]. O , F[illegible] 5. MDD [illegible]. T[illegible] 100 [illegible]

[illegible] 52. H[illegible] ([illegible]), [illegible] MLP [illegible]. C[illegible]

A[illegible] ACC [illegible]. I[illegible] LC, MS [illegible]. O[illegible] ACC, AUC [illegible]. T[illegible]

F[illegible]. T[illegible] 28 [illegible] 14 [illegible]. T[illegible]. F[illegible] CMS-N [illegible].

Key Points

- [illegible] DL [illegible] CMS-N [illegible] LC, MS [illegible]. I[illegible] LC, MS [illegible].
- [illegible] CMS-N [illegible].
- B[illegible] CMS-N [illegible].

Supplementary data

S[illegible] Briefings in Bioinformatics [illegible].

C[illegible]: N[illegible].

Funding

T[illegible] D[illegible] I[illegible] C[illegible] P[illegible] I[illegible] P[illegible] (202207I); P[illegible] L[illegible] P[illegible] D[illegible] S[illegible] T[illegible] (2021 JH2/10200004); N[illegible] S[illegible] F[illegible] L[illegible] P[illegible], C[illegible] (2024-MS-175); D[illegible] S[illegible] T[illegible] T[illegible] I[illegible] S[illegible] P[illegible] (2022RG17); [illegible] 2023 B[illegible] S[illegible] R[illegible] P[illegible] I[illegible] P[illegible] D[illegible] E[illegible] (J[illegible] T[illegible] D2023050).

Data availability

T[illegible]. T[illegible] CMS-N [illegible] ://[illegible]/CMS-N.

References

1. H[illegible] S, K[illegible] I, D[illegible] L. et al. S[illegible]? Int J Mol Sci 2022;**24**:578. [illegible]://[illegible]/10.3390/[illegible]24010578.

2. J, R QN, J K. C . *Diagnostics (Basel)* 2022;**13**:2. ://. /10.3390/ 13010002.

3. E AH. T NMR - . *Methods Mol Biol* 2015;**1277**:161 93. ://. /10.1007/978-1-4939-2377-9 13.

4. R S, H L, B B. *et al.* S LC-MS- . *J Mass Spectrom Adv Clin Lab* 2023;**28**:47 55. ://. /10.1016/ . .2023.02.003.

5. Q S, C , H. *et al.* S : . *Signal Transduct Target Ther* 2023;**8**:132. ://. /10.1038/ 41392-023-01399-3.

6. R FH, S R, B K. *et al.* A . *Heliyon* 2023;**9**: 18909. ://. /10.1016/ .2023. 18909.

7. L D, R B, N DR. *et al.* B - 1, : . *Transl Psychiatry* 2018;**8**:10. ://. /10.1038/ 41398-017-0056-8.

8. P DI, S J, R G. *et al.* B /I G N- . *Sci Rep* 2018;**8**:179. ://. /10.1038/ 41598-017-17500-0.

9. S RA, B N, J MP. *et al.* V . *Sci Rep* 2021;**11**:1406. ://. /10.1038/ 41598-020-80873-2.

10. A L, R B, E O. *et al.* A : . *Brain Behav Immun* 2023;**114**: 242 54. ://. /10.1016/ . .2023.08.025.

11. D L, M V, V AK. C . *ACS Omega* 2020;**5**:28877 88. ://. /10.1021/ .0 04568.

12. , L , L D. *et al.* P . *Proteomics Clin Appl* 2021;**15**: 2000058. ://. /10.1002/ .202000058.

13. R N, P G, F DA. *et al.* P - . *Sci Rep* 2023;**13**:6332. ://. /10.1038/ 41598-023-33359- .

14. C J, M M, M R. *et al.* O - . *Bioinformatics* 2021;**37**: 245 53. ://. /10.1093/ / 311.

15. L M, M , H *et al.* R - . *Molecules* 2023;**28**:7430. ://. /10.3390/ 2821 7430.

16. M M, K C, F. *et al.* A . *Cell Syst* 2021;**12**:759 70. ://. /10.1016/ . .2021.06.006.

17. C K, D -S E, C S. *et al.* A : . *IEEE J Biomed Health Inform* 2023;**27**:457 68. ://. /10.1109/JBHI.2022. 3217015.

18. K G, M -L A, D D. D . *Neuropsychopharmacology* 2021;**46**:176 90. ://. /10.1038/ 41386-020-0767- .

19. D H, L , F. *et al.* A - - - MS . *Proteomics* 2020;**20**: 1900344. ://. /10.1002/ .201900344.

20. S , S S, . *et al.* A - . *Cell Discov* 2022;**8**:85. ://. /10.1038/ 41421-022-00442- .

21. , L J, G M. *et al.* R - - . *Adv Sci (Weinh)* 2020;**7**:2002021. ://. /10.1002/ .202002021.

22. J, C N, . *et al.* C - . *Bioinformatics* 2023;**39**:39. ://. /10.1093/ / 422.

23. C J, L CH, EG. *et al.* D - - : . *Mol Syst Biol* 2021;**17**: 10240. ://. /10.15252/ .202110240.

24. R SJ, S D, S D. *et al.* L - - . *Transl Psychiatry* 2023;**13**:44. ://. /10.1038/ 41398-023-02321-9.

25. U S, K A, H ME. *et al.* C . *BMC Med Inform Decis Mak* 2019;**19**:281. ://. /10.1186/ 12911-019-1004-8.

26. N LD, S N, R GPS. A IL-5 - - . *Comput Biol Med* 2023;**158**:106864. ://. /10.1016/ . .2023.106864.

27. I T, G JL, R AJ. S : . *Behav Ther* 2020;**51**:675 87. ://. /10.1016/ . .2020.05.002.

28. H GS, K HK, C RV. *et al.* D : . *J Intern Med* 2018;**284**: 603 19. ://. /10.1111/ .12822.

29. M P, S U, P CM. *et al.* A , - , . *Ann Card Anaesth* 2019;**22**:407 11. ://. /10.4103/ .ACA 94 19.

30. I , L , F M. *et al.* G - . *MethodsX* 2023;**10**:102009. ://. /10.1016/ . .2023.102009.

31. J B , T D, CG. S - . *JAMIA Open* 2023;**6**: 086. ://. /10.1093/ / 086.

32. L S, , F . A - - : 7- - . *BMC Psychiatry* 2022;**22**:816. ://. /10.1186/ 12888-022-04439-4.

33. C T, H K, H S. *et al.* D - AI . *Nat Commun* 2024;**15**:524. ://. /10.1038/ 41467-023-43095-4.

34. O D, A D, H A. *et al.* E - . *JMIR AI* 2023;**2**: 41205. ://. /10.2196/41205.

35. R MJ, K MA, K SH. *et al.* D GB

[illegible] SHAP. *Sci Rep* 2023;**13**:6263. [illegible]://[illegible]/10.1038/[illegible]41598-023-33525-0.

36. D[illegible] Q, M[illegible] JG. P[illegible] SHAP (P SHAP) [illegible]. *PLoS Comput Biol* 2022;**18**:[illegible]1009736. [illegible]://[illegible]/10.1371/[illegible].1009736.
37. T[illegible] P, C[illegible] T, [illegible] CH. *et al.* P[illegible]. *Crit Care* 2020;**24**:478. [illegible]://[illegible]/10.1186/[illegible]13054-020-03179-9.
38. M[illegible], [illegible] B, [illegible] F. *et al.* P[illegible]-[illegible]. *Nat Commun* 2021;**12**:3008. [illegible]://[illegible]/10.1038/[illegible]41467-021-22756-2.
39. A[illegible] A, T[illegible] L, S[illegible] O. *et al.* P[illegible]: [illegible]. *Bioinformatics* 2010;**26**:1340[illegible]7. [illegible]://[illegible]/10.1093/[illegible]/[illegible]134.
40. G[illegible], [illegible] J, G[illegible] N. *et al.* A PO[illegible]. *Int J Mol Sci* 2013;**14**:22132[illegible]48. [illegible]://[illegible]/10.3390/[illegible]141122132.
41. L[illegible], [illegible] S. K[illegible]-[illegible]. *Entropy (Basel)* 2023;**25**:325. [illegible]://[illegible]/10.3390/[illegible]25020325.
42. K[illegible] D, S[illegible] A, J[illegible] S. *et al.* D[illegible]. *J Psychiatr Res* 2013;**47**:162[illegible]7. [illegible]://[illegible]/10.1016/[illegible].2012.09.017.
43. H[illegible] CSH, S[illegible] M[illegible] T, T[illegible] G[illegible] N. T[illegible] RNA [illegible]. *Compr Psychiatry* 2023;**121**:152363. [illegible]://[illegible]/10.1016/[illegible].2022.152363.
44. C[illegible] BP, [illegible] R[illegible] NS. B[illegible]-[illegible] (BDNF) [illegible]. *Prog Neuropsychopharmacol Biol Psychiatry* 2021;**108**:110151. [illegible]://[illegible]/10.1016/[illegible].2020.110151.
45. [illegible] C, M[illegible], [illegible] L. *et al.* A[illegible]. *J Zhejiang Univ Sci* 2021;**48**:18[illegible]29.
46. B[illegible] H, P[illegible] D, B[illegible] P. T[illegible]. *Data Min Knowl Discov* 2023;**38**:2596[illegible]632. [illegible]://[illegible]/10.1007/[illegible]10618-023-00924-[illegible].
47. B[illegible] AC, [illegible] M[illegible] F, R[illegible]-C[illegible] TA. *et al.* D[illegible] PTSD [illegible] ICD-11 PTSD [illegible]. *Depress Anxiety* 2019;**36**:490[illegible]8. [illegible]://[illegible]/10.1002/[illegible].22881.
48. K[illegible] P, [illegible] H, R[illegible] PP. *et al.* C[illegible]-[illegible] EEG [illegible]-[illegible]. *Sci Rep* 2023;**13**:7467. [illegible]://[illegible]/10.1038/[illegible]41598-023-34298-2.
49. C[illegible] VA, B[illegible] BH, K[illegible] S. *et al.* S[illegible]: [illegible]. *Neuropsychopharmacology* 2019;**44**:1224[illegible]32. [illegible]://[illegible]/10.1038/[illegible]41386-019-0369-9.
50. H[illegible], [illegible] C, F[illegible] *et al.* C[illegible] C[illegible]: [illegible]-[illegible]. *J Affect Disord* 2022;**319**:221[illegible]8. [illegible]://[illegible]/10.1016/[illegible].2022.08.123.
51. L[illegible], [illegible], M[illegible] L. *et al.* I[illegible]. *J Nurs Sci* 2020;**35**:85[illegible]8.
52. [illegible], G[illegible], R[illegible] *et al.* A[illegible]. *J Lanzhou Univ Med Sci* 2022;**48**:5[illegible]10.

Briefings in Bioinformatics, 2025, **26(1)**, [illegible]554
[illegible]://[illegible]/10.1093/[illegible]/[illegible]554
P[illegible] S[illegible] P[illegible]