UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

BIBLIOGRAPHY

(2017–2020). Continuous automated model evaluation. https://www.cameo3d.org/sp/6-months/. Accessed: 2022-02-03.

(2017–2020). Protein structure prediction center. https://predictioncenter.org/. Accessed: 2022-02-03.

(2018). Neuron cell. https://www.cusabio.com/Cell-Marker/Neuron-Cell.html.

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A. I., Cowie, A., Figurnov, M., Fuchs, F. B., Gladman, H., Jain, R., Khan, Y. A., Low, C. M. R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E. D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, *630*, 493–500.

AlGhamdi, R., Aziz, A., Alshehri, M., Pardasani, K. R., & Aziz, T. (2021). Deep learning model with ensemble techniques to compute the secondary structure of proteins. *The Journal of Supercomputing*, *77*, 5104–5119.

AlQuraishi, M. (2019). ProteinNet: a standardized data set for machine learning of protein structure. *BMC Bioinformatics*, *20*, 311.

Altschul, S. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*, 3389–3402.

ao, N. P., Heinrich, J., Stolte, C., Sabir, K. S., Buckley, M. J., Tabor, B., Signal, B., Gloss, B. S., Hammang, C. J., Rost, B., Schafferhans, A., & O'Donoghue, S. I. (2015). Unexpected features of the dark proteome. *Proceedings of the National Academy of Sciences*, *112*, 15898–15903.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom, M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Aydin, Z., Azginoglu, N., Bilgin, H. I., & Celik, M. (2019). Developing structural profile matrices for protein secondary structure and solvent accessibility prediction. *Bioinformatics*, *35*, 4004–4010.

Batuwita, R., & Palade, V. (2012). Adjusted geometric-mean: a novel performance measure for imbalanced bioinformatics datasets learning. *Journal of bioinformatics and computational biology*, *10*, 1250003.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2005). Genbank. *Nucleic Acids Research*, *33*, D34–D38.

Birzele, F., & Kramer, S. (2006). A new representation for protein secondary structure prediction based on frequent patterns. *Bioinformatics*, *22*, 2628–2634.

Bohr, H., Bohr, J., Brunak, S., Cotterill, R. M., Lautrup, B., Nørskov, L., Olsen, O. H., & Petersen, S. B. (1988). Protein secondary structure and homology by neural networks the α-helices in rhodopsin. *FEBS Letters*, *241*, 223–228.

de Brevern, A. G. (2020). Impact of protein dynamics on secondary structure prediction. *Biochimie*, *179*, 14–22.

Bruce Alberts, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2004). *Molecular Biology of the Cell, 6th edn*. Garland Science.

Callaway, E. (2022). What's next for AlphaFold and the AI protein-folding revolution. *Nature*, *604*, 234–238.

Carkli Yavuz, B., Yurtay, N., & Ozkan, O. (2018). Prediction of protein secondary structure with clonal selection algorithm and multilayer perceptron. *IEEE Access*, *6*, 45256–45261.

Chang, D. T.-H., Huang, H.-Y., Syu, Y.-T., & Wu, C.-P. (2008). Real value prediction of protein solvent accessibility using enhanced pssm features. *BMC Bioinformatics*, *9*, S12.

Chen, Y. (2017). Long sequence feature extraction based on deep learning neural network for protein secondary structure prediction. In *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)* (pp. 843–847). IEEE volume 2017-Janua.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Cheng, J., Liu, Y., & Ma, Y. (2020). Protein secondary structure prediction based on integration of CNN and LSTM model. *Journal of Visual Communication and Image Representation*, *71*, 102844.

Chou, P. Y., & Fasman, G. D. (1974). Prediction of protein conformation. *Biochemistry*, *13*, 222–245.

Chowdhury, R., Bouatta, N., Biswas, S., Rochereau, C., Church, G. M., Sorger, P. K., & AlQuraishi, M. (2021). Single-sequence protein structure prediction using language models from deep learning. *bioRxiv*, .

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, .

Clark, D. P., Pazdernik, N. J., & McGehee, M. R. (2019). Chapter 14 - protein structure and function. In D. P. Clark, N. J. Pazdernik, & M. R. McGehee (Eds.), *Molecular Biology (Third Edition)* (pp. 445–483). Academic Cell. (Third edition ed.).

Clark, K., Luong, M., Le, Q. V., & Manning, C. D. (2020). ELECTRA: pre-training text encoders as discriminators rather than generators. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020* (pp. 1–18). OpenReview.net.

Cole, C., Barber, J. D., & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Research*, *36*, W197–W201.

Consortium, T. U. (2007). The Universal Protein Resource (UniProt). *Nucleic Acids Research*, *36*, D190–D195.

Cuff, J. A., & Barton, G. J. (1999). Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins: Structure, Function, and Bioinformatics*, *34*, 508–519.

Cuff, J. A., & Barton, G. J. (2000). Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins: Structure, Function, and Bioinformatics*, *40*, 502–511.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Dai, L., & Zhou, Y. (2011). Characterizing the existing and potential structural space of proteins by large-scale multiple loop permutations. *Journal of Molecular Biology*, *408*, 585–595.

Dai, Z., Yang, Z., Yang, Y., Carbonell, J. G., Le, Q. V., & Salakhutdinov, R. (2019). Transformer-XL: Attentive language models beyond a fixed-length context. In A. Korhonen, D. R. Traum, & L. Màrquez (Eds.), *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers* (pp. 2978–2988). Association for Computational Linguistics.

Dencelin, X. L., & Ramkumar, T. (2018). An approach to predict protein secondary structure using deep learning in spark based big data computing framework. In *2018 4th International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)* (pp. 25–30). IEEE.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics.

Dionysiou, A., Agathocleous, M., & B, C. C. (2018). Convolutional neural networks in combination with support vector machines for complex sequential data classification. In V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, & I. Maglogiannis (Eds.), *International Conference on Artificial Neural Networks (ICANN)* (pp. 444–455). Cham: Springer International Publishing volume 11140 of *Lecture Notes in Computer Science*.

Dongardive, J., & Abraham, S. (2017). Reaching optimized parameter set: protein secondary structure prediction using neural network. *Neural Computing and Applications*, *28*, 1947–1974.

Dor, O., & Zhou, Y. (2006). Achieving 80% ten-fold cross-validated accuracy for secondary structure prediction by large-scale training. *Proteins: Structure, Function, and Bioinformatics*, *66*, 838–845.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Drori, I., Dwivedi, I., Shrestha, P., Wan, J., Wang, Y., He, Y., Mazza, A., Krogh-Freeman, H., Leggas, D., Sandridge, K., Nan, L., Thakoor, K., Joshi, C., Goenka, S., Keasar, C., & Pe'er, I. (2018). High quality protein Q8 secondary structure prediction by diverse neural network architectures. *arXiv*, (pp. 1–10).

Dubchak, I., Muchnik, I., & Kim, S.-H. (1997). Protein folding class predictor for scop: approach based on global descriptors. In *Proceedings. International Conference on Intelligent Systems for Molecular Biology* (pp. 104–107). volume 5.

Elnaggar, A., Heinzinger, M., Dallago, C., Rihawi, G., Wang, Y., Jones, L., Gibbs, T., Feher, T., Angerer, C., Steinegger, M. et al. (2020). Prottrans: towards cracking the language of life's code through self-supervised deep learning and high performance computing. *arXiv preprint arXiv:2007.06225*, .

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J., & Hassabis, D. (2022). Protein complex prediction with alphafold-multimer. *bioRxiv*, .

Fang, C., Li, Z., Xu, D., & Shang, Y. (2019). MUFold-SSW: a new web server for predicting protein secondary structures, torsion angles and turns. *Bioinformatics*, *36*, 1293–1295.

Fang, C., Shang, Y., & Xu, D. (2017). A new deep neighbor residual network for protein secondary structure prediction. In *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 66–71). IEEE.

Fang, C., Shang, Y., & Xu, D. (2018). MUFOLD-SS: New deep inception-inside-inception networks for protein secondary structure prediction. *Proteins: Structure, Function and Bioinformatics*, *86*, 592–598.

Faraggi, E., Yang, Y., Zhang, S., & Zhou, Y. (2009). Predicting continuous local structure and the effect of its substitution for secondary structure in fragment-free protein structure prediction. *Structure*, *17*, 1515–1527.

Fidelis, K., Rost, B., & Zemla, A. (1999). A modified definition of SOV, a segment-based measure. *Proteins*, *223*, 220–223.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Finkelstein, A. V., & Ptitsyn, O. B. (1971). Statistical analysis of the correlation among amino acid residues in helical, β-stractural and non-regular regions of globular proteins. *Journal of Molecular Biology*, *62*, 613–624.

Firat, O., Cho, K., & Bengio, Y. (2016). Multi-way, multilingual neural machine translation with a shared attention mechanism. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 866–875). Stroudsburg, PA, USA: Association for Computational Linguistics.

Fischer, D., & Eisenberg, D. (1996). Protein fold recognition using sequence-derived predictions. *Protein Science*, *5*, 947–955.

Fox, N. K., Brenner, S. E., & Chandonia, J. M. (2014). SCOPe: Structural classification of proteins - extended, integrating SCOP and ASTRAL data and classification of new structures. *Nucleic Acids Research*, *42*, 304–309.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *Journal of Molecular Biology*, *120*, 97–120.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*.

Görmez, Y., & Aydın, Z. (2022). IGPRED-MultiTask: A deep learning model to predict protein secondary structure, torsion angles and solvent accessibility. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, (pp. 1–12).

Görmez, Y., & Aydın, Z. (2021). *ROSE: A Novel Approach for Protein Secondary Structure Prediction* volume 76. Springer International Publishing.

Görmez, Y., Sabzekar, M., & Aydın, Z. (2021). IGPRED: Combination of convolutional neural and graph convolutional networks for protein secondary structure prediction. *Proteins: Structure, Function, and Bioinformatics*, *89*, 1277–1288.

Gromiha, M. M. (2010). Chapter 1 - proteins. In M. M. Gromiha (Ed.), *Protein Bioinformatics* (pp. 1–27). Singapore: Academic Press.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Guo, Y., Li, W., Wang, B., Liu, H., & Zhou, D. (2019). DeepACLSTM: Deep asymmetric convolutional long short-term memory neural models for protein secondary structure prediction. *BMC Bioinformatics*, *20*, 1–12.

Guo, Y., Wang, B., Li, W., & Yang, B. (2018). Protein secondary structure prediction improved by recurrent neural networks integrated with two-dimensional convolutional neural networks. *Journal of Bioinformatics and Computational Biology*, *16*, 1850021.

Guo, Y., Wu, J., Ma, H., Wang, S., & Huang, J. (2020a). Bagging MSA learning: Enhancing low-quality PSSM with deep learning for accurate protein structure property prediction. In *Research in Computational Molecular Biology: 24th Annual International Conference, RECOMB 2020, Padua, Italy, May 10–13, 2020, Proceedings 24* (pp. 88–103). Springer.

Guo, Y., Wu, J., Ma, H., Wang, S., & Huang, J. (2020b). Protein ensemble learning with atrous spatial pyramid networks for secondary structure prediction. In *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 17–22). IEEE.

Guo, Z., Hou, J., & Cheng, J. (2021). DNSS2: Improved ab initio protein secondary structure prediction using advanced deep learning architectures. *Proteins: Structure, Function, and Bioinformatics*, *89*, 207–217.

Hanson, J., Paliwal, K., Litfin, T., Yang, Y., & Zhou, Y. (2018). Accurate prediction of protein contact maps by coupling residual two-dimensional bidirectional long short-term memory with convolutional neural networks. *Bioinformatics*, *34*, 4039–4045.

Hanson, J., Paliwal, K., Litfin, T., Yang, Y., & Zhou, Y. (2019). Improving prediction of protein secondary structure, backbone angles, solvent accessibility and contact numbers by using predicted contact maps and an ensemble of recurrent and residual convolutional neural networks. *Bioinformatics*, *35*, 2403–2410.

Hasic, H., Buza, E., & Akagic, A. (2017). A hybrid method for prediction of protein secondary structure based on multiple artificial neural networks, . (pp. 1195–1200).

Hattori, L. T., Vargas Benitez, C. M., & Lopes, H. S. (2017). A deep bidirectional long short-term memory approach applied to the protein secondary structure prediction

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

problem. In *2017 IEEE Latin American Conference on Computational Intelligence (LA-CCI)* (pp. 1–6). IEEE.

Heffernan, R., Paliwal, K., Lyons, J., Dehzangi, A., Sharma, A., Wang, J., Sattar, A., Yang, Y., & Zhou, Y. (2015). Improving prediction of secondary structure, local backbone angles and solvent accessible surface area of proteins by iterative deep learning. *Scientific Reports*, *5*, 11476.

Heffernan, R., Paliwal, K., Lyons, J., Singh, J., Yang, Y., & Zhou, Y. (2018). Single-sequence-based prediction of protein secondary structures and solvent accessibility by deep whole-sequence learning. *Journal of Computational Chemistry*, *39*, 2210–2216.

Heffernan, R., Yang, Y., Paliwal, K., & Zhou, Y. (2017). Capturing non-local interactions by long short-term memory bidirectional recurrent neural networks for improving prediction of protein secondary structure, backbone angles, contact numbers and solvent accessibility. *Bioinformatics*, *33*, 2842–2849.

Heinzinger, M., Elnaggar, A., Wang, Y., Dallago, C., Nechaev, D., Matthes, F., & Rost, B. (2019). Modeling aspects of the language of life through transfer-learning protein sequences. *BMC Bioinformatics*, *20*, 723.

Ho, C.-T., Huang, Y.-W., Chen, T.-R., Lo, C.-H., & Lo, W.-C. (2021). Discovering the ultimate limits of protein secondary structure prediction. *Biomolecules*, *11*.

Hobohm, U., & Sander, C. (1994). Enlarged representative set of protein structures. *Protein Science*, *3*, 522–524.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*, 1735–1780.

Høie, M. H., Kiehl, E. N., Petersen, B., Nielsen, M., Winther, O., Nielsen, H., Hallgren, J., & Marcatili, P. (2022). NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. *Nucleic Acids Research*, . Gkac439.

Howard Holley, l., & Karplus, M. (1989). Protein secondary structure prediction with a neural network. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 152–156.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Hu, H., Li, Z., Elofsson, A., & Xie, S. (2019a). A Bi-LSTM based ensemble algorithm for prediction of protein secondary structure. *Applied Sciences*, *9*, 3538.

Hu, Y., Nie, T., Shen, D., & Yu, G. (2019b). A deep aggregated model for protein secondary structure prediction. *International Journal of Data Mining and Bioinformatics*, *22*, 231.

Ibraheem, R., Connor, K., & Salome, B. (2024). Biochemistry, tertiary protein structure, .

Jalal, S. I., Zhong, J., & Kumar, S. (2019). Protein secondary structure prediction using multi-input convolutional neural network. In *2019 SoutheastCon* (pp. 1–5). IEEE.

Jana, N. D., Das, S., & Sil, J. (2018). Backgrounds on protein structure prediction and metaheuristics, . (pp. 1–28).

Jing, X., Dong, Q., Hong, D., & Lu, R. (2020). Amino acid encoding methods for protein sequences: A comprehensive review and assessment. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *17*, 1918–1931.

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, *292*, 195–202.

Ju, F., Zhu, J., Zhang, Q., Wei, G., Sun, S., Zheng, W.-M., & Bu, D. (2021). SeqSetNet: directly exploiting multiple sequence alignment for protein secondary structure prediction. *Bioinformatics*, *38*, 990–996.

Juan, S.-H., Chen, T.-R., & Lo, W.-C. (2020). A simple strategy to enhance the speed of protein secondary structure prediction without sacrificing accuracy. *PLOS ONE*, *15*, e0235153.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., & Hassabis, D. (2021). Highly accurate protein structure prediction with alphafold. *Nature*, *596*, 583–589.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, *22*, 2577–2637.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., & Phillips, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*, *181*, 662–666.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, .

Klausen, M. S., Jespersen, M. C., Nielsen, H., Jensen, K. K., Jurtz, V. I., Sønderby, C. K., Sommer, M. O. A., Winther, O., Nielsen, M., Petersen, B., & Marcatili, P. (2019). NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. *Proteins: Structure, Function, and Bioinformatics*, *87*, 520–527.

Kotowski, K., Smolarczyk, T., Roterman-Konieczna, I., & Stapor, K. (2021). ProteinUnet- an efficient alternative to SPIDER3-single for sequence-based prediction of protein secondary structures. *Journal of Computational Chemistry*, *42*, 50–59.

Kumar, P., Bankapur, S., & Patil, N. (2020). An enhanced protein secondary structure prediction using deep learning framework on hybrid profile based features. *Applied Soft Computing Journal*, *86*, 105926.

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A lite BERT for self-supervised learning of language representations. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020* (pp. 1–17). OpenReview.net.

Le, Q., Sievers, F., & Higgins, D. G. (2017). Protein multiple sequence alignment benchmarking through secondary structure prediction. *Bioinformatics*, *33*, btw840.

Levin, J. M., Pascarella, S., Argos, P., & Garnier, J. (1993). Quantification of secondary structure prediction improvement using multiple alignments. *Protein Engineering, Design and Selection*, *6*, 849–854.

Levitan, I. B., & Kaczmarek, L. K. (2015). *The neuron: cell and molecular biology*. Oxford University Press, USA.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Li, B., Mendenhall, J., Capra, J. A., & Meiler, J. (2021). A multitask deep-learning method for predicting membrane associations and secondary structures of proteins. *Journal of Proteome Research*, *20*, 4089–4100.

Li, Z., & Yu, Y. (2016). Protein secondary structure prediction using cascaded convolutional and recurrent neural networks. *IJCAI International Joint Conference on Artificial Intelligence*, *2016-Janua*, 2560–2567.

Lin, Z., Lanchantin, J., & Qi, Y. (2016). MUST-CNN: A multilayer shift-and-stitch deep convolutional architecture for sequence-based protein structure prediction. *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, (pp. 27–34).

Liu, T., & Wang, Z. (2018). SOV-refine: A further refined definition of segment overlap score and its significance for protein structure similarity. *Source Code for Biology and Medicine*, *13*, 1–10.

Liu, Y., Chen, Y., & Cheng, J. (2016). Feature extraction of protein secondary structure using 2D convolutional neural network. In *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (pp. 1771–1775). IEEE.

Liu, Y., Cheng, J., Ma, Y., & Chen, Y. (2018). Protein secondary structure prediction based on two dimensional deep convolutional neural networks. *2017 3rd IEEE International Conference on Computer and Communications, ICCC 2017*, *2018-Janua*, 1995–1999.

Liu, Z., Gong, Y., Bao, Y., Guo, Y., Wang, H., & Lin, G. N. (2021). TMPSS: A deep learning-based predictor for secondary structure and topology structure prediction of alpha-helical transmembrane proteins. *Frontiers in Bioengineering and Biotechnology*, *8*, 1–12.

Lo Conte, L. (2000). SCOP: a structural classification of proteins database. *Nucleic Acids Research*, *28*, 257–259.

Lomize, M. A., Lomize, A. L., Pogozheva, I. D., & Mosberg, H. I. (2006). OPM: Orientations of proteins in membranes database. *Bioinformatics*, *22*, 623–625.

Long, S., & Tian, P. (2019). Protein secondary structure prediction with context convolutional neural network. *RSC Advances*, *9*, 38391–38396.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom, M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Luong, T., Pham, H., & Manning, C. D. (2015). ffective approaches to attention-based neural machine translation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 1412–1421). Stroudsburg, PA, USA: Association for Computational Linguistics.

Lyu, Z., Wang, Z., Luo, F., Shuai, J., & Huang, Y. (2021). Protein secondary structure prediction with a reductive deep learning method. *Frontiers in Bioengineering and Biotechnology*, *9*, 1–8.

Maas, A. L., Hannun, A. Y., Ng, A. Y. et al. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of the international conference on machine learning (ICML)* (p. 3). Atlanta, GA volume 30.

Magnan, C. N., & Baldi, P. (2014). SSpro / ACCpro 5: almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles , machine learning and structural similarity. *Bioinformatics*, *30*, 2592–2597.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta*, *405*, 442–451.

McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, *16*, 404–405.

Meiler, J., Müller, M., Zeidler, A., & Schmäschke, F. (2001). Generation and evaluation of dimension-reduced amino acid parameter representations by artificial neural networks. *Molecular modeling annual*, *7*, 360–369.

Midic, U., Dunker, A. K., & Obradovic, Z. (2007). Exploring alternative knowledge representations for protein secondary-structure prediction. *International Journal of Data Mining and Bioinformatics*, *1*, 286.

Mitchell, E. M., Artymiuk, P. J., Rice, D. W., & Willett, P. (1990). Use of techniques derived from graph theory to compare secondary structure motifs in proteins. *Journal of Molecular Biology*, *212*, 151–166.

Moffat, L., & Jones, D. T. (2021). Increasing the accuracy of single sequence prediction methods using a deep semi-supervised learning framework. *Bioinformatics*, .

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom, M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Momen-Roknabadi, A., Sadeghi, M., Pezeshk, H., & Marashi, S.-A. (2008). Impact of residue accessible surface area on the prediction of protein secondary structures. *BMC Bioinformatics*, *9*, 357.

Mufassirin, M. M., Newton, M. H., Rahman, J., & Sattar, A. (2023). Multi-s3p: Protein secondary structure prediction with specialized multi-network and self-attention-based deep learning model. *IEEE Access*, .

Müller, B., Reinhardt, J., & Strickland, M. T. (2012). *Neural networks: an introduction*. Springer Science & Business Media.

Naderi-Manesh, H., Sadeghi, M., Arab, S., & Moosavi Movahedi, A. A. (2001). Prediction of protein surface accessibility with information theory. *Proteins: Structure, Function, and Bioinformatics*, *42*, 452–459.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)* (pp. 807–814).

Nelson, D. L., Lehninger, A. L., & Cox, M. M. (2008). *Lehninger principles of biochemistry*. Macmillan.

de Oliveira, G. B., Pedrini, H., & Dias, Z. (2020). Ensemble of bidirectional recurrent networks and random forests for protein secondary structure prediction. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)* (pp. 311–316). IEEE volume 2020-July.

Oliveira, G. B. D., Pedrini, H., & Dias, Z. (2020). Fusion of BLAST and ensemble of classifiers for protein secondary structure prediction. In *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)* (pp. 308–315). IEEE.

Ozkan, S. B., Wu, G. A., Chodera, J. D., & Dill, K. A. (2007). Protein folding by zipping and assembly. *Proceedings of the National Academy of Sciences*, *104*, 11987–11992.

Pauling, L., Corey, R. B., & Branson, H. R. (1951). The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences*, *37*, 205–211.

PDB (2023). PDB Statistics: Overall growth of released structures per year.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Penel, S., Morrison, R., Dobson, P. D., Mortishire-Smith, R. J., & Doig, A. J. (2003). Length preferences and periodicity in -strands. antiparallel edge -sheets are more likely to finish in non-hydrogen bonded rings. *Protein Engineering Design and Selection*, *16*, 957–961.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)* (pp. 2227–2237). New Orleans, Louisiana: Association for Computational Linguistics.

Pollastri, G., Przybylski, D., Rost, B., & Baldi, P. (2002). Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins: Structure, Function, and Bioinformatics*, *47*, 228–235.

Qi, Y., Oja, M., Weston, J., & Noble, W. S. (2012). A unified multitask architecture for predicting local protein properties. *PLoS ONE*, *7*, e32235.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*, 1–67.

Rashid, S., Saraswathi, S., Kloczkowski, A., Sundaram, S., & Kolinski, A. (2016). Protein secondary structure prediction using a small training set (compact model) combined with a complex-valued neural network approach. *BMC Bioinformatics*, *17*, 362.

Remmert, M., Biegert, A., Hauser, A., & Söding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature Methods*, *9*, 173–175.

Richards, F. M., & Kundrot, C. E. (1988). Identification of structural motifs from protein coordinate data: Secondary structure and first-level supersecondary structure*. *Proteins: Structure, Function, and Bioinformatics*, *3*, 71–84.

Rives, A., Meier, J., Sercu, T., Goyal, S., Lin, Z., Liu, J., Guo, D., Ott, M., Zitnick, C. L., Ma, J., & Fergus, R. (2021). Biological structure and function emerge from

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, *118*, e2016239118.

Rizqiana, A., & Afiahayati, A. (2023). N-grams modeling for protein secondary structure prediction: Exploring local features and optimal cnn parameters. In *BIO Web of Conferences* (p. 01008). EDP Sciences volume 75.

Rohl, C. A., Strauss, C. E., Misura, K. M., & Baker, D. (2004). Protein structure prediction using Rosetta. In *Methods in Enzimology* (pp. 66–93). volume 383.

Rost, B. (1996). [31] PHD: Predicting one-dimensional protein structure by profile-based neural networks. In *Computer Methods for Macromolecular Sequence Analysis* (pp. 525–539). Academic Press volume 266 of *Methods in Enzymology*.

Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Engineering, Design and Selection*, *12*, 85–94.

Rost, B. (2001). Review: Protein secondary structure prediction continues to rise. *Journal of Structural Biology*, *134*, 204–218.

Rost, B., & Eyrich, V. A. (2001). Eva: Large-scale analysis of secondary structure prediction. *Proteins: Structure, Function, and Bioinformatics*, *45*, 192–199.

Rost, B., & Sander, C. (1983). Prediction of protein secondary structure at better than 70% accuracy. *Journal of Molecular Biology*, *232*, 584–599.

Rost, B., Sander, C., & Schneider, R. (1994). Redefining the goals of protein secondary structure prediction. *Journal of Molecular Biology*, *235*, 13–26.

Sander, C., & Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins: Structure, Function, and Bioinformatics*, *9*, 56–68.

Sauder, J. M., Arthur, J. W., & Dunbrack Jr., R. L. (2000). Large-scale comparison of protein sequence alignment algorithms with structure alignments. *Proteins: Structure, Function, and Bioinformatics*, *40*, 6–22.

Scheraga, H. A. (1960). Structural studies of ribonuclease. iii. a model for the secondary and tertiary structure. *Journal of the American Chemical Society*, *82*, 3847–3852.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Shapovalov, M., Dunbrack, R. L., & Vucetic, S. (2020). Multifaceted analysis of training and testing convolutional neural networks for protein secondary structure prediction. *PLOS ONE*, (pp. 1–38).

Shuai-yan, Z., & Yi-hui, L. (2017). A novel radical group encoding method for protein secondary structure prediction. In *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(* (pp. 939–943). IEEE.

Shuai-yan, Z., Yi-hui, L., & Jin-yong, C. (2017). The prediction of protein secondary structure based on auto encoder. In *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* 61375013 (pp. 2065–2069). IEEE.

Sidi, T., & Keasar, C. (2020). Redundancy-weighting the PDB for detailed secondary structure prediction using deep-learning models. *Bioinformatics*, *36*, 3733–3738.

Sievers, F., & Higgins, D. G. (2019). QuanTest2: benchmarking multiple sequence alignments using secondary structure prediction. *Bioinformatics*, *36*, 90–95.

Sillitoe, I., Lewis, T. E., Cuff, A., Das, S., Ashford, P., Dawson, N. L., Furnham, N., Laskowski, R. A., Lee, D., Lees, J. G., Lehtinen, S., Studer, R. A., Thornton, J., & Orengo, C. A. (2015). CATH: Comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Research*, *43*, D376–D381.

Singh, J., Litfin, T., Paliwal, K., Singh, J., Hanumanthappa, A. K., & Zhou, Y. (2021). SPOT-1D-Single: improving the single-sequence-based prediction of protein secondary structure, backbone angles, solvent accessibility and half-sphere exposures using a large training set and ensembled deep learning. *Bioinformatics*, (pp. 1–9).

Singh, J., Litfin, T., Singh, J., Paliwal, K., & Zhou, Y. (2022a). SPOT-Contact-LM: improving single-sequence-based prediction of protein contact map using a transformer language model. *Bioinformatics*, *38*, 1888–1894.

Singh, J., Paliwal, K., Litfin, T., Singh, J., & Zhou, Y. (2022b). Reaching alignment-profile-based accuracy in predicting protein secondary and tertiary structural properties without alignment. *Scientific Reports*, *12*, 7607.

Skolnick, J., Kihara, D., & Zhang, Y. (2004). Development and large scale benchmark testing of the PROSPECTOR_3 threading algorithm. *Proteins: Structure, Function, and Bioinformatics*, *56*, 502–518.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Stapor, K., Kotowski, K., Smolarczyk, T., & Roterman, I. (2022). Lightweight proteinunet2 network for protein secondary structure prediction: a step towards proper evaluation. *BMC Bioinformatics*, *23*, 100.

Steinegger, M., Mirdita, M., & Söding, J. (2019). Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods*, *16*, 603–606.

Steinegger, M., & Söding, J. (2017). Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology*, *35*, 1026–1028.

Stormo, G. D. (2009). An introduction to sequence similarity ("homology") searching. *Current Protocols in Bioinformatics*, *27*, 3.1.1–3.1.7.

Sutanto, V., Sukma, Z., & Afiahayati, A. (2021). Predicting secondary structure of protein using hybrid of convolutional neural network and support vector machine. *International Journal of Intelligent Engineering and Systems*, *14*, 232–243.

Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R., & Wu, C. H. (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, *23*, 1282–1288.

Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., Wu, C. H., & the UniProt Consortium (2014). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics*, *31*, 926–932.

Tai, C.-H., Bai, H., Taylor, T. J., & Lee, B. (2014). Assessment of template-free modeling in casp10 and roll. *Proteins: Structure, Function, and Bioinformatics*, *82*, 57–83.

Torrisi, M., Kaleel, M., & Pollastri, G. (2018). Porter 5: Fast, state-of-the-art ab initio prediction of protein secondary structure in 3 and 8 classes. *bioRxiv*, (pp. 1–7).

Torrisi, M., Kaleel, M., & Pollastri, G. (2019). Deeper profiles and cascaded recurrent and convolutional neural networks for state-of-the-art protein secondary structure prediction. *Scientific Reports*, *9*, 12374.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Uddin, M. R., Mahbub, S., Saifur Rahman, M., & Bayzid, M. S. (2020). SAINT: Self-attention augmented inception-inside-inception network improves protein secondary structure prediction. *Bioinformatics*, *36*, 4599–4608.

UniProt (2023). UniProtKB/TrEMBL protein database release 2021_03 statistics.

Urban, G., Torrisi, M., Magnan, C. N., Pollastri, G., & Baldi, P. (2020). Protein profiles: Biases and protocols. *Computational and Structural Biotechnology Journal*, *18*, 2281–2289.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *2017-Decem*, 5999–6009.

Walz, J., & Rahal, I. (2018). Secondary protein structure prediction combining protein structural class, relative surface accessibility, and contact number. *International Journal of Data Science*, *3*, 68.

Wang, G., & Dunbrack, J., Roland L. (2003). PISCES: a protein sequence culling server. *Bioinformatics*, *19*, 1589–1591.

Wang, J., Cheng, J., Zhao, Z., & Lu, W. (2019). Protein secondary structure prediction using ensemble of LSTM neural networks. In *2019 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE)* (pp. 241–244). IEEE.

Wang, L., & Cheng, J. (2018). Protein secondary structure prediction using autoencoder network and bayes classifier. *IOP Conference Series: Materials Science and Engineering*, *322*, 062008.

Wang, Q., Wang, B., Xu, Z., Wu, J., Zhao, P., Li, Z., Wang, S., Huang, J., & Cui, S. (2021). PSSM-Distil: Protein secondary structure prediction (PSSP) on low-quality PSSM by knowledge distillation with contrastive learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*, 617–625.

Wang, Q., Wei, J., Zhou, Y., Lin, M., Ren, R., Wang, S., Cui, S., & Li, Z. (2022). Prior knowledge facilitates low homologous protein secondary structure prediction with DSM distillation. *Bioinformatics*, .

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Wang, S., Li, W., Liu, S., & Xu, J. (2016a). RaptorX-Property: a web server for protein structure property prediction. *Nucleic Acids Research*, (pp. 1–6).

Wang, S., Peng, J., Ma, J., & Xu, J. (2016b). Protein secondary structure prediction using deep convolutional neural fields. *Scientific Reports*, *6*, 18962.

Wang, Y., Mao, H., & Yi, Z. (2017). Protein secondary structure prediction by using deep learning method. *Knowledge-Based Systems*, *118*, 115–123.

Weissenow, K., Heinzinger, M., & Rost, B. (2022). Protein language-model embeddings for fast, accurate, and alignment-free protein structure prediction. *Structure*, *30*, 1169–1177.e4.

Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. *PhD thesis, Committee on Applied Mathematics, Harvard University, Cambridge, MA*, .

Woo, S.-K., Park, C.-B., & Lee, S.-W. (2005). Protein secondary structure prediction using sequence profile and conserved domain profile. In *International Conference on Intelligent Computing* (pp. 1–10). Springer.

Wood, M. J., & Hirst, J. D. (2005). Protein secondary structure prediction with dihedral angles. *Proteins: Structure, Function, and Bioinformatics*, *59*, 476–481.

Wu, S., Skolnick, J., & Zhang, Y. (2007). Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biology*, *5*, 17.

Xu, G., Wang, Q., & Ma, J. (2020). OPUS-TASS: A protein backbone torsion angles and secondary structure predictor based on ensemble neural networks. *Bioinformatics*, *36*, 5021–5026.

Xu, Y., & Cheng, J. (2020). Protein secondary structure prediction using CNN and random forest. In *Artificial Intelligence and Security. ICAIS 2020. Communications in Computer and Information Science* (pp. 267–277).

Yang, L., Wei, P., Zhong, C., Li, X., & Tang, Y. Y. (2020). Protein structure prediction based on BN-GRU method. *International Journal of Wavelets, Multiresolution and Information Processing*, .

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Yang, W., Hu, Z., Zhou, L., & Jin, Y. (2022a). Protein secondary structure prediction using a lightweight convolutional network and label distribution aware margin loss. *Knowledge-Based Systems*, *237*, 107771.

Yang, W., Liu, Y., & Xiao, C. (2022b). Deep metric learning for accurate protein secondary structure prediction. *Knowledge-Based Systems*, *242*, 108356.

Yang, Y., Gao, J., Wang, J., Heffernan, R., Hanson, J., Paliwal, K., & Zhou, Y. (2016). Sixty-five years of the long march in protein secondary structure prediction: the final stretch? *Briefings in Bioinformatics*, *19*, 482–494.

Yang, Y., Heffernan, R., Paliwal, K., Lyons, J., Dehzangi, A., Sharma, A., Wang, J., Sattar, A., & Zhou, Y. (2017). SPIDER2: A package to predict secondary structure, accessible surface area, and main-chain torsional angles by deep neural networks. In *Methods in Molecular Biology* (pp. 55–63). volume 1484.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (pp. 1–11). Curran Associates, Inc. volume 32.

Yi, P., Albert Y, Z., & Series Editors (2010). *Introduction to Protein Structure Prediction*. John Wiley & Sons, Inc.

Yi, T.-M., & Lander, E. S. (1993). Protein secondary structure prediction using nearest-neighbor methods. *Journal of Molecular Biology*, *232*, 1117–1129.

Zamani, M., & Kremer, S. C. (2016). Protein secondary structure prediction through a novel framework of secondary structure transition sites and new encoding schemes. In *2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (pp. 1–7). IEEE.

Zemla, A., Venclovas, v., Fidelis, K., & Rost, B. (1999). A modified definition of sov, a segment-based measure for protein secondary structure prediction assessment. *Proteins: Structure, Function, and Bioinformatics*, *34*, 220–223.

Zhang, B., Li, J., & Lü, Q. (2018). Prediction of 8-state protein secondary structures by a novel deep learning architecture. *BMC Bioinformatics*, *19*, 293.

UNIVERSITAS
GADJAH MADA
SELF-ATTENTION AND BIDIRECTIONAL GATED RECURRENT UNIT FOR PROTEIN SECONDARY STRUCTURE PREDICTION
DEWI PRAMUDI ISMI, Prof. Dr.-Ing. Mhd. Reza M. I. Pulungan, S.Si., M.Sc.; Afiahayati, S.Kom., M.Cs., Ph.D.
Universitas Gadjah Mada, 2024 | Diunduh dari http://etd.repository.ugm.ac.id/

Zhao, Y., & Liu, Y. (2021). OCLSTM: Optimized convolutional and long short-term memory neural network model for protein secondary structure prediction. *PLOS ONE*, *16*, e0245982.

Zhao, Y., Zhang, H., & Liu, Y. (2020). Protein secondary structure prediction based on generative confrontation and convolutional neural network. *IEEE Access*, *8*, 199171–199178.

Zhou, J., & Troyanskaya, O. (2014). Deep supervised and convolutional generative stochastic network for protein secondary structure prediction. In *International conference on machine learning* (pp. 745–753). PMLR.

Zhou, J., Wang, H., Zhao, Z., Xu, R., & Lu, Q. (2018). CNNH_PSS: Protein 8-class secondary structure prediction by convolutional neural network with highway. *BMC Bioinformatics*, *19*.

Zhou, Y., Duan, Y., Yang, Y., Faraggi, E., & Lei, H. (2011). Trends in template/fragment-free protein structure prediction. *Theoretical Chemistry Accounts*, *128*, 3–16.

Zhou, Y., & Karplus, M. (1999). Interpreting the folding kinetics of helical proteins. *Nature*, *401*, 400–403.