Identifying Parkinson’s Disease from Typing Behavior with BERT

Scott Thompson, Cynthia Xu
2023 Fall 266 NLP

Image source

A neurodegenerative illness we can help manage symptoms for

10 Million people impacted globally

Potential to improve quality of life

Impacts speech and typing

ML for scalable screening

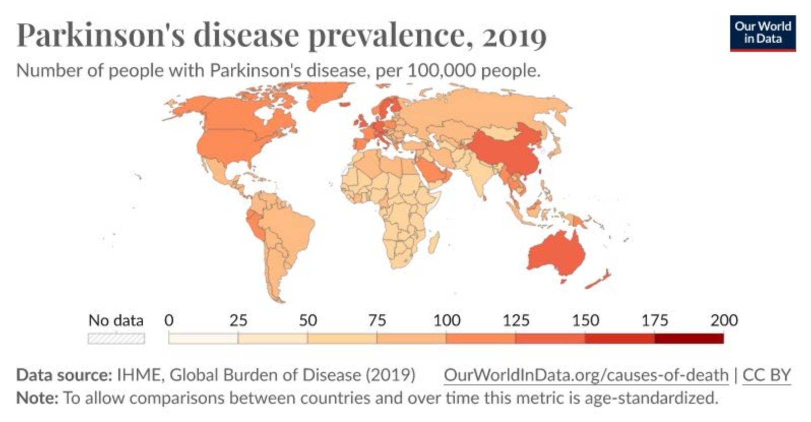


Image source

Prior research leverages speech and typing

Vocal features and transcribed speech

Typing behavior, with an emphasis on keystroke timing

RNN, LSTM, CNN, BERT, BETO, etc. all utilized

A copy-typing dataset with individual keystrokes

Kindly provided by Dhir, et al (2020)

1470 responses from PD patients and 1919 responses from healthy controls (HC)

Individuals copy-typed 10-15 word long sentences from Wikipedia articles

Data representation - characters, words and flight time

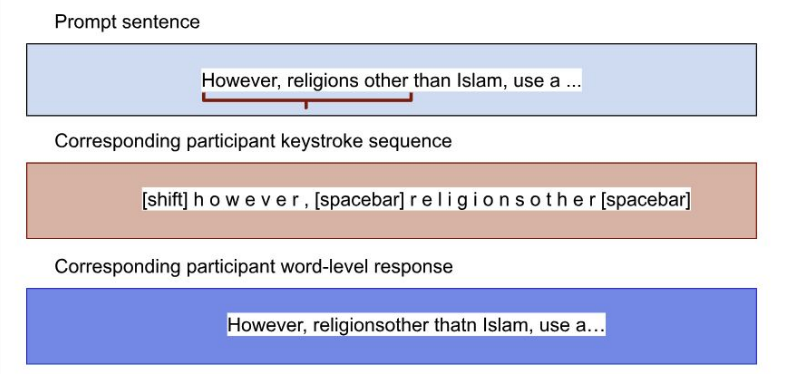


Figure 1. An example of the typing data.

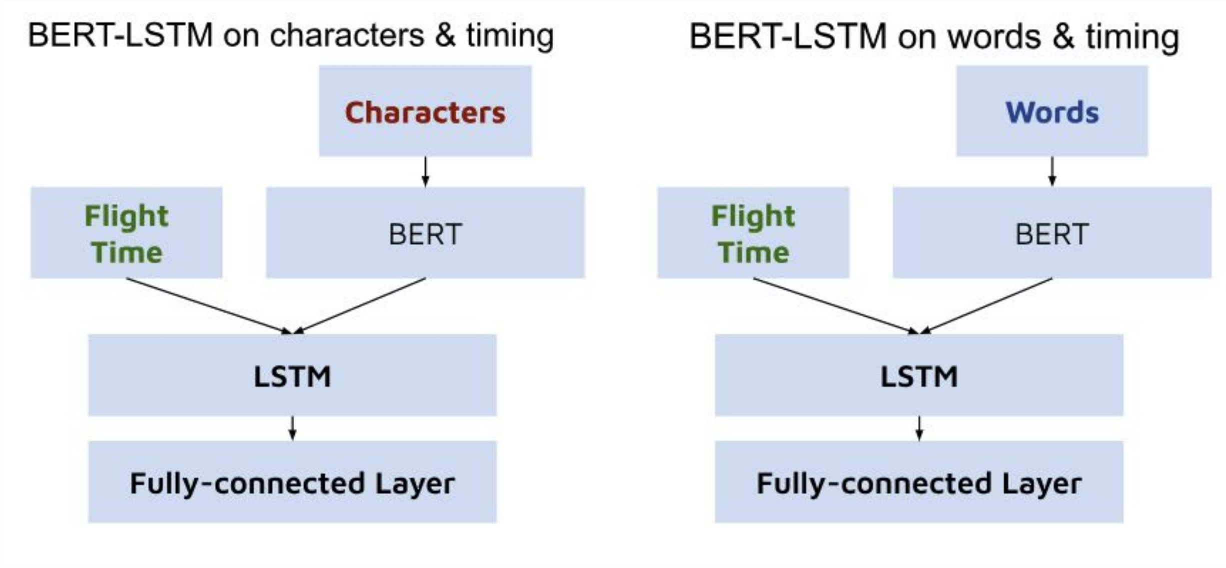
Key “flight time” measuring the time between releasing one key and pressing another

A range of different model inputs and architectures

Model Name	BERT on characters	BERT-LSTM on characters	BERT-LSTM on characters & flight time	BERT on words	BERT-LSTM on words	BERT-LSTM on words & flight time	BERT-LSTM on words, characters & flight time	CNN-LSTM on characters & flight time
Data Input	- Characters	- Characters	- Characters - Flight Time	- Words	- Words	- Words - Flight Time	- Characters - Words - Flight Time	- Characters - Flight Time
Embedding	BERT							CBOW
Additional Layer	None	LSTM	LSTM	None	LSTM	LSTM	LSTM	CNN LSTM

Table 1. Model variations for experimental setup.

BERT-LSTM architecture



Strong performance as measured by AUC

Model	Accuracy	Precision	Recall	F1	AUC
Baseline	0.484 (0.041)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.500 (0.000)
BERT on characters	0.533 (0.000)	0.533 (0.000)	1.000 (0.000)	0.695 (0.000)	0.482 (0.054)
BERT-LSTM on characters	0.489 (0.038)	0.178 (0.308)	0.333 (0.577)	0.232 (0.401)	0.515 (0.005)
BERT-LSTM on characters & flight time	0.700 (0.017)	0.940 (0.011)	0.467 (0.027)	0.623 (0.027)	**0.846 (0.008)**
BERT on words	0.489 (0.038)	0.178 (0.308)	0.333 (0.577)	0.232 (0.401)	0.502 (0.035)
BERT-LSTM on words	0.487 (0.024)	0.516 (0.027)	0.400 (0.205)	0.431 (0.160)	0.512 (0.025)
BERT-LSTM on words & flight time	0.669 (0.015)	0.906 (0.021)	0.423 (0.024)	0.576 (0.025)	**0.838 (0.006)**
BERT-LSTM on words, characters, & flight time	0.704 (0.023)	0.934 (0.032)	0.479 (0.032)	0.633 (0.033)	**0.845 (0.004)**
CNN-LSTM on characters & flight time	0.832 (0.010)	0.919 (0.016)	0.751 (0.010)	0.827 (0.011)	**0.848 (0.006)**

Attention is not all you need… but flight time is

No semantic meaning for BERT to understand → Poor BERT performance

Flight time dramatically improves AUC

Words are no better than characters

Telling right from wrong - Classification examples

True Negative Example

Key flight times: 188.0, 141.0, 124.0, 157.0, 219.0, 125.0, 125…
Tokens: [shift] t h e y [spacebar] f o u g h t [spacebar] a [spacebar] t h i r t y…

False Negative Example

Key flight times: 153.0, 467.0, 70.0, 227.0, 95.0, 236.0, 139.0...
Tokens: [shift] t h e y [spacebar] f o u g h [spacebar] a [spacebar] t h i r t y…

True Positive Example

Key flight times: 128.685, 773.28, 2813.75, 168.015, 188.895, 6...
Tokens: s[shift] p l i t - f i n g e r [spacebar] a i m i n g [spacebar] r e…

False Positive Example

Key flight times: 87.79992, 227.60023, 88.29992, 188.79984, 247...
Tokens: [shift] h o w e v e r, [unidentified] [spacebar] t h e r e [spacebar] i s…

Figure 4. Classification examples BERT-LSTM on characters & flight time model

Conclusion

Achieved benchmark AUC

However, to achieve a new SOTA we suggest patient-generated text

Semantic differences that allow attention-based strategies to shine

References

- Neil Dhir, Mathias Edman, Álvaro Sanchez Ferro, Tom Stafford, and Colin Bannard. 2020. Identifying robust markers of Parkinson's disease in typing behaviour using a CNN-LSTM network. In Proceedings of the 24th Conference on Computational Natural Language Learning, pages 578–595, Online. Association for Computational Linguistics.
- Parkinson's Foundation. Accessed 2 Oct 2023. Statistics - Get informed about Parkinson's disease with these key numbers. Online.
- S. Tripathi, T. Arroyo-Gallego and L. Giancardo. Keystroke-Dynamics for Parkinson's Disease Signs Detection in an At-Home Uncontrolled Population: A New Benchmark and Method, in IEEE Transactions on Biomedical Engineering, vol. 70, no. 1, pp. 182-192, Jan. 2023, doi: 10.1109/TBME.2022.3187309.
- Nijhawan R, Kumar M, Arya S, Mendirtta N, Kumar S, Towfek SK, Khafaga DS, Alkahtani HK, Abdelhamid AA. A Novel Artificial-Intelligence-Based Approach for Classification of Parkinson's Disease Using Complex and Large Vocal Features. Biomimetics. 2023; 8(4):351.
- Escobar-Grisales D, Ríos-Urrego CD, Orozco-Arroyave JR. Deep Learning and Artificial Intelligence Applied to Model Speech and Language in Parkinson's Disease. Diagnostics. 2023; 13(13):2163.
- Chintalapudi N, Battineni G, Hossain MA, Amenta F. Cascaded Deep Learning Frameworks in Contribution to the Detection of Parkinson's Disease. *Bioengineering (Basel)*. 2022 Mar 12;9(3):116. doi: 10.3390/bioengineering9030116. PMID: 35324805; PMCID: PMC8945200.
- Wentao Ma, Yiming Cui, Chenglei Si, Ting Liu, Shijin Wang, and Guoping Hu. 2020. CharBERT: Character-aware Pre-trained Language Model. In Proceedings of the 28th International Conference on Computational Linguistics, pages 39–50, Barcelona, Spain (Online). International Committee on Computational Linguistics.

References

- Katsunori Yokoi, Yurie Iribe, Norihide Kitaoka, Takashi Tsuboi, Keita Hiraga, Yuki Satake, Makoto Hattori, Yasuhiro Tanaka, Maki Sato, Akihiro Hori, Masahisa Katsuno. Analysis of spontaneous speech in Parkinson's disease by natural language processing, Parkinsonism & Related Disorders, Volume 113, 2023, 105411, ISSN 1353-8020
- Vásquez-Correa, J.C. et al. (2019). Convolutional Neural Networks and a Transfer Learning Strategy to Classify Parkinson's Disease from Speech in Three Different Languages. In: Nyström, I., Hernández Heredia, Y., Milián Núñez, V. (eds) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2019. Lecture Notes in Computer Science(), vol 11896. Springer, Cham.
- Pfeiffer, H. C. V., Løkkegaard, A., Zoetmulder, M., Friberg, L., & Werdelin, L. (2014). Cognitive impairment in early-stage non-demented Parkinson's disease patients. Acta Neurologica Scandinavica, 129(5), 307-318.
- Soumen Roy, Utpal Roy, Devadatta Sinha, Rajat Kumar Pal. (2023). Imbalanced ensemble learning in determining Parkinson's disease using Keystroke dynamics. In: Expert Systems with Applications, Volume 217, 2023, 119522, ISSN 0957-4174.

Additional sources/contributions: Advice/guidance from Jennifer Zhu and other instructors, course materials (including async lectures, live sessions and provided notebooks), assignments, ChatGPT (coding support), StackOverflow (coding support) and Medium (coding support).