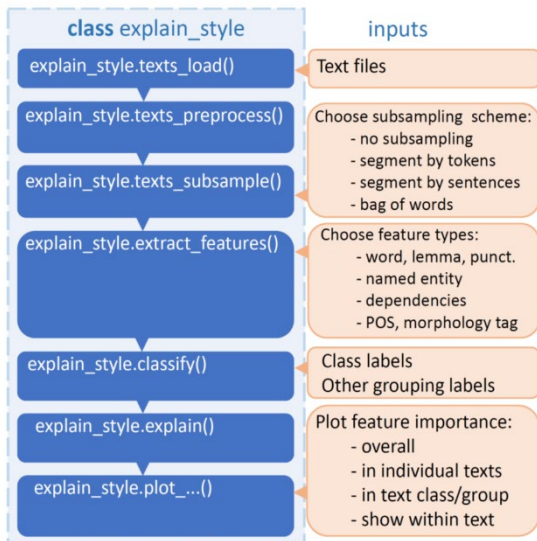


Implementing interpretable models in stylometric analysis

JK Ochab Jagiellonian University in Kraków
T Walkowiak Wrocław University of Science and Technology
jeremi.ochab@uj.edu.pl, tomasz.walkowiak@pwr.edu.pl



Package structure

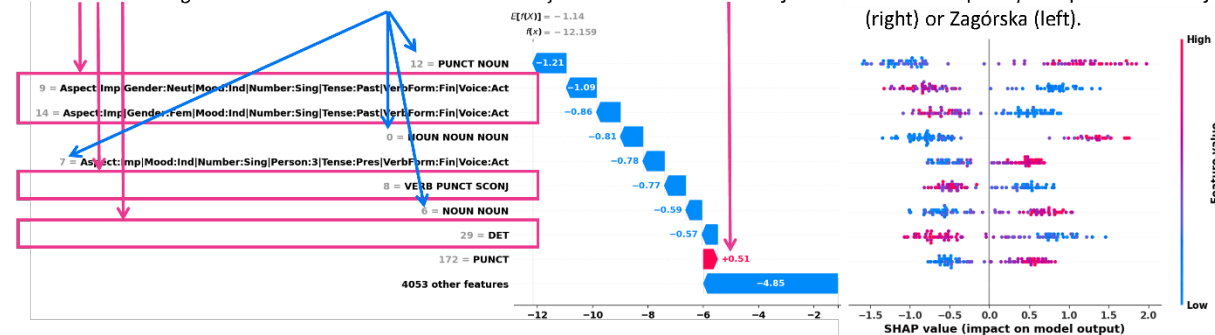


High incidence of features characteristic of Zagórska.

Absence/low incidence of features characteristic of Dukaj.

High incidence of a feature characteristic of Dukaj.

The colour indicates whether **high/low** feature frequency is expected of Dukaj (right) or Zagórska (left).



MAIN AIM

Provide linguistically explainable features that lead to a given text classification.

Features characteristically overrepresented in Zagórska (negative SHAP value).

najlepiej – powtórzyłem . I może tak **było** istotnie . Lecz z **każdym** wymówionym **słowem** **pokój** stawał się

mroczniejszy ; tylko jej czoło , gładkie i białe , **jaśniało** niewyczerpanym **światłem** **wiary** i miłości . – Pan był jego

przyjacielem – **ciągnęła** . – **Przyjacielem** – powtórzyła trochę głośniejsz . – Pan musiał być jego przyjacielem , jeśli

panu to **powierzył** , jeśli pana do mnie przysłał . **Czuję** , że mogę z panem mówić – i – och , muszę z panem

Eder, M., Kestemont, M., Rybicki, J. (2016). The R Journal, 8 (1), 1–15.
Ke, G., Meng, Q., Finley T., et al. (2017). Advances in neural information processing systems, 30, 3146–3154.
Lundberg, S. M., Erion, G., Chen, H., et al. (2020). Nature Machine Intelligence, 2 (1), 2522–5839.
Montani, I., Honnibal, M., Honnibal, et al. (2023). “explosion/spaCy: v3.7.2: Fixes for APIs and requirements.”
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Journal of Machine Learning Research, 12, 2825–2830

Research financed by the European Regional Development Fund as a part of the 2014–2020 Smart Growth Operational Programme, CLARIN—Common Language Resources and Technology Infrastructure, project no. POIR.04.02.00-00C002/19. This venture has been made possible thanks to the funding provided by the Jagiellonian University’s Flagship Project “Digital Humanities Lab”.



Republic of Poland

European Union
European Regional Development Fund

