12th ICAMDATA - 25-29 September 2022, Mola di Bari (Italy)
International Conference on Atomic and Molecular Data and Their Applications
Book of Abstracts

Machine learning for prediction of Stark broadening parameters

A. Zakuskin [1], S.M. Zaytsev [1], T.A. Labutin[1]

[1] Department of Chemistry, Lomonosov Moscow State University, 1 Leninskiye Gory, bld. 3, Moscow, Russia, 119234

Besides the transition probabilities the Stark broadening parameters are of a key importance for modelling the spectra of various emission sources. For example, lack of Stark parameters for atomic lines observed in spectra of white dwarfs prevents the correct analysis and interpretation of their spectra. Although a huge number of experimental and theoretical studies for retrieving Stark parameters have been being done for decades and new data are regularly added, only a few per cent of known transitions are covered by existing databases. Getting new values is still very important and often non-trivial task. Experimental studies are strongly limited by inhomogeneity of plasma sources, spectral interferences, lack of spectral resolution and non-equilibrium interactions between particles. Broadening of some emission lines, such as doubly charged atomic ions, cannot be observed in laboratory plasma due to low values of Stark broadening parameters. On the other hand, quantum chemical calculations limited by high computational costs, and, in many cases, by strong interaction of electronic states or relativistic effects for heavy atoms. At the same time, rapid development of machine and deep learning methods resulted in their high efficiency and accuracy of prediction in many scientific areas including spectroscopy. Artificial neural networks (ANNs) have been successfully implemented for various tasks from analysis of hyperspectral images to prediction of reaction rates, equilibrium constants, chemical and physical properties of the substance by its structure. All these considerations led us to the idea of application of machine learning (ML) for prediction of Stark parameters.

We designed a table format for representation of electronic levels configuration, term and energy for each atomic or ionic transition. The database consists of ≈1000 transitions with experimental values of Stark parameters only. Availability of parameters measured at different temperatures allows addition of temperature parameter into the database and prediction of temperature-dependent values of Stark parameters by implemented ML methods.

Several models of classical machine learning (KNN, Random Forest, Gradient Boosting etc.) as well as ANNs (TabNet and custom NN) were trained and tested for the prediction of Stark broadening and shift parameters with independent subsets. The highest accuracy was demonstrated by boosting algorithms, but all the models have insufficient accuracy in prediction of anomalously large values of parameters. It is important that ANNs are not competitive with classical ML models yet, but their accuracy strongly depends on the size of the training dataset, so filling up the database with new experimental values would possibly result in further accuracy improvement. We also reviewed models' ability for generalization by testing them on the independent subset of transitions for 3 chemical elements that were not seen by the models in the training set.

Predicted Stark parameters for large number of transitions were added to the thermodynamic model simulating low-temperature plasma emission. Comparison of experimental and simulated spectra with predicted Stark parameters clearly shows that knowledge of parameters for large number of transitions brings benefits for spectra interpretation and plasma diagnostics.

The work was supported by the Fellowship from Non-commercial Foundation for the Advancement of Science and Education INTELLECT.