Publ. Astron. Obs. Belgrade No. 102 (2022), 243 - 246 *Contributed Paper*

MODELING OF STARK SPECTRAL LINE BROADENING BY MACHINE LEARNING ALGORITHMS

IRINEL TAPALAGA[1], IVAN TRAPARIĆ[2], NORA TRKLJA BOCA[1], JAGOŠ PURIĆ[1] and IVAN P. DOJČINOVIĆ[1]

[1]*University of Belgrade, Faculty of Physics, P. O. Box 44, 11000 Belgrade, Serbia*

[2]*University of Belgrade, Institute of Physics, Pregrevica 118, 11080 Belgrade, Serbia*

E-mail nora@ff.bg.ac.rs

Abstract. Stark broadening caused by free electrons in plasma and its dependence on atomic parameters available in NIST and Stark-b atomic databases has been investigated using a new method based on the machine learning (ML) approach. The correlation parameter obtained by artificial intelligence (AI) is slightly better than the one obtained by classical methods, but the scope of application is much wider. ML algorithms successfully identified quantum nature by analyzing atomic parameters. The biggest issue of classical analysis, which is infinite spectral line broadening for high ionization stages, was resolved by AI with a saturation tendency.

1. INTRODUCTION AND THEORETICAL BACKGROUND

Stark broadening of spectral lines of neutral atoms and ions is used in science for a number of problems in various physical conditions (see Tapalaga et al. 2022). Recent research indicates the importance and usefulness of searching for possible types of regularities in the framework of a Stark broadening investigation. Still, existing tables with calculated and measured Stark widths have a big lack of data. There is a need for Stark widths data in the wide range of chemical elements, plasma temperature and electron densities. In this paper a correlation between Stark broadening and environment parameters, such as the ionization potential of the upper level of the corresponding transition, electron density and temperature, will be investigated using modern ML algorithms. If this method proves to be accurate enough, the process of calculating the value of Stark widths will be significantly accelerated and facilitated.

The general formula for Stark width calculation in the impact approximation (see Griem 1974) is very complicated, it cannot be resolved exactly, so it is useful

to use different approaches in the calculation. The regularity approach which correlates Stark width of spectral line (ω) expressed in [rad/s], electron density (N_e), electron temperature (T_e) and positive value of electron binding energy on the upper level of the transition (χ), expressed in [eV], is given by Puric and Scepanovic 1999. (Eq. 1)

$$\omega = Z_e^k a N_e f(T_e) \chi^{-b} \tag{1}$$

where Z_e = 1, 2, 3... for neutrals, singly charged ions, ... respectively and it represents the rest core charge of the ionized emitter and a, b and k are coefficients independent of electron concentration and ionization potential for a particular transition and the rest core charge of the emitter.

It is expected that spectral series within an isoelectronic sequence show regularity behavior because a wide range of atomic/ionic parameters depend on the electron number. In the last decade we have investigated Stark broadening regularities within spectral series of individual elements, individual isoelectronic sequences and within two spectral series of isoelectronic sequences simultaneously. The present investigation goes one step further and analyses all elements for which there are available data needed for Stark broadening investigation, simultaneously, using machine learning approach. The aim is to find the best possible model which correlates Stark width of spectral line with all available parameters for transition of interest (atomic parameters and environmental parameters).

2. DATASET CREATION AND DATA CLEANING

In order to create dataset, two public repositories connected with atomic spectroscopy are used. First one is Stark B database, see Sahal-Brechot et al. 2020, where the parameters of Stark broadening for different emitters are given. The features taken from this database are: chemical element, ionization stage, upper and lower level of spectral transitions, Stark broadening, the environment temperature and electron density in environment. For analysis purpose, Stark widths expressed in angstroms are converted in radian per second.

To ensure better results we enriched features taken from Stark B database with ones taken from NIST Atomic Spectra database (see Kramida et al. 2019): binding energy of both upper and lower transition levels, ground level energy, total angular momentum quantum number (*J)* of both upper and lower transition level, as well as principal (*n)* and orbital (*l)* quantum numbers and ionization energy. The algorithm for connecting those two databases to form our own works is described below. For every transition connected with certain chemical element, we take the electronic configuration of both upper and lower levels from Stark B database. Then we look for that particular element in NIST database and compare the electronic configurations. If they match, then we take the binding energy of those levels, their

principal quantum number, orbital quantum number and total angular momentum quantum number and finally the ionization energy of that atom.
After data cleaning, dataset contains 53 emitters and 34973 spectral lines and follows a normal distribution.

3. MODEL CREATION AND TRAINING

For model creation and training, public Python package Sci-kit learn is used. Four models have been created, every being Pipeline with two steps. In each object of Pipeline class, the first step was data scaling using StandardScaler, and in second step we made our predictions with defined model. Considered models were: Linear Regression, Decision Tree Regressor, Random Forest Regressor and Gradient Boosting Regressor. The dataset was split into training and test dataset using train test split method, leaving 25% of the data for testing. To rank the performance of models, we used best Coefficient of Determination, R^2, value obtained after GridSearchCV algorithm finished. As a result, we got that the best R^2 value was for Random Forest Regressor having $R^2 = 0.95$ for n estimators = 100. Random Forest is a learning method that operates by constructing a large number of decision trees during the training process, see Tapalaga et al. 2022. It is simple to use and shows high performance for a wide variety of tasks, making it one of the most popular ML algorithms in different sciences. Random forests are an effective tool in predicting new data, in our case new atomic parameters.

4. RESULTS

The Random Forest model is used to calculate Stark broadening data for spectral series within neutral lithium Li I, see Tapalaga et al. 2022. Fig 1. shows the dependence of the Stark width (ω) on the reciprocal value of the electron binding energy at the upper level of the transition (χ^{-1}) for 2s-np and 3s-np transitions within lithium atom. Calculated Stark widths (red lines) for transitions within analyzed series are represented with existing known values of Stark widths data at the same graphs (see Fig. 1). The functional dependence obtained using the ML algorithm describes the quantum structure of the energy levels of lithium atoms. From the model lines (red lines) it can be concluded that the model successfully (within the error) indicates the quantum nature of atomic transitions and that other results do not make physical sense, but only jumps.

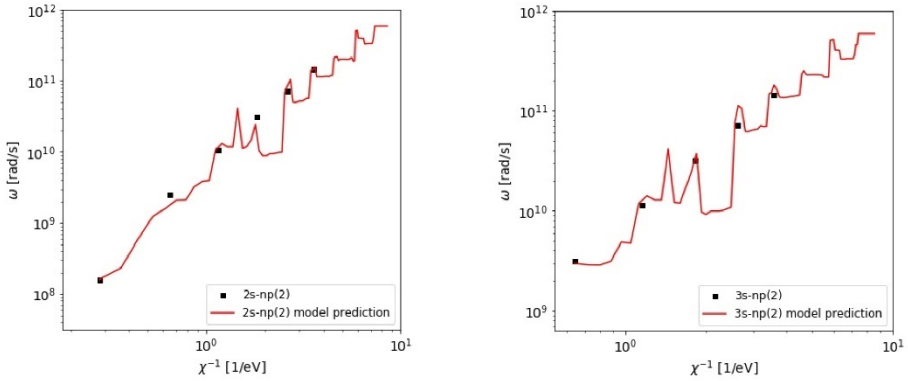

Figure 1: Stark widths regularities within 2s-np and 3s-np spectral series of Li I ($T = 30000$ K, $N_e = 10^{20}$ m^{-3})

5. CONCLUSION

Analysis of spectral data on Stark broadening for 53 different emitters and 34973 lines by ML algorithms was done with more success than it was previously done by classical methods of data analysis. Random forest has scored an average of $R^2 = 0.95$ which makes it an excellent choice for Stark broadening calculations. With standard known methods for Stark width calculation, it is not possible to calculate Stark widths for levels for which energy values of the closest perturbing levels are missing, but ML algorithms enable calculation in these situations, too.

References

Griem, H.R. : 1974, *Spectral Line Broadening by Plasmas*, Academic Press, New York

Kramida A, Ralchenko Yu, Reader J and NIST ASD Team (2019). NIST Atomic Spectra Database (version 5.7.1), [Online]. Available: https://physics.nist.gov/asd [Tue Mar 24, 2020]. National Institute of Standards and Technology, Gaithersburg, MD.

Purić, J., Šćepanović, M. : 1999, *Astrophys. J*, **521**, 490.

Sahal-Brechot S., Dimitrijević M. S., Moreau N. (2020) STARK-B database, [online]. Available: http://stark-b.obspm.fr [February 20, 2020]. Observatory of Paris, LERMA and Astronomical Observatory of Belgrade

Tapalaga, I., Traparić, I., Trklja Boca N., Purić, J., Dojčinović, P. I. : 2022, *Neural Computing and Application*, **34**, 6349