

Revisión de algunos algoritmos de Machine Learning aplicados a la clasificación de espectros estelares

Content

La astronomía moderna enfrenta el reto de analizar grandes volúmenes de datos generados por misiones como Gaia, particularmente en su tercera entrega (DR3), que incluye espectros de baja resolución para más de 219 millones de objetos. En este contexto, el presente trabajo realiza una revisión sistemática de diversos algoritmos de aprendizaje automático (Machine Learning) aplicados a la clasificación de espectros estelares. Se analizan técnicas como Random Forest, Support Vector Machine, Artificial Neural Networks, Gradient Boosting y Naive Bayes, evaluando su desempeño mediante métricas como precisión, sensibilidad, F1-score y Kappa de Cohen. Asimismo, se discuten enfoques modernos como el aprendizaje activo y comparaciones con métodos tradicionales basados en inferencia bayesiana. Los resultados de los trabajos analizados indican que los algoritmos basados en Artificial Neural Networks y Gradient Boosting ofrecen mayor precisión, especialmente con conjuntos de datos balanceados. Esta revisión establece las bases metodológicas para una propuesta de investigación orientada a mejorar la precisión y eficiencia de la clasificación espectral mediante técnicas de aprendizaje profundo aplicadas a los datos de Gaia DR3.

Palabras Clave: clasificación espectral, aprendizaje automático, Gaia DR3, astronomía computacional, algoritmos.

Summary

Modern astronomy faces the challenge of analyzing large volumes of data generated by missions such as Gaia, particularly in its third release (DR3), which includes low-resolution spectra for more than 219 million objects. In this context, the present work systematically reviews various machine learning algorithms applied to the classification of stellar spectra. Techniques such as Random Forest, Support Vector Machine, Artificial Neural Networks, Gradient Boosting, and Naive Bayes are analyzed, evaluating their performance using metrics such as precision, recall, F1-score, and Cohen's Kappa. Moreover, modern approaches like active learning and comparisons with traditional methods based on Bayesian inference are discussed. The results from the reviewed studies indicate that algorithms based on Artificial Neural Networks and Gradient Boosting offer greater accuracy, especially with balanced datasets. This review establishes the methodological foundation for a research proposal aimed at improving the accuracy and efficiency of spectral classification through deep learning techniques applied to Gaia DR3 data.

Keywords: classification of stellar spectra, machine learning, Gaia DR3, computational astronomy, algorithms.

Primary author(s) : Mrs. PEREZ-APARICIO, ARELI (Universidad de Guadalajara)

Co-author(s) : Dr. NAVARRO, SILVANA (UNIVERSIDAD DE GUADALAJARA); Dr. CORRAL, LUIS JOSE (UNIVERSIDAD DE GUADALAJARA)

Presenter(s) : Mrs. PEREZ-APARICIO, ARELI (Universidad de Guadalajara)