A MACHINE LEARNING APPROACH TO CHARACTERIZING THE CHEMICAL INVENTORY OF ORION-KL

HALEY N. SCOLATI, JOSHUA CARDER, *Department of Chemistry, University of Virginia, Charlottesville, VA, USA*; ERIC HERBST, *Department of Chemistry, The University of Virginia, Charlottesville, VA, USA*; BRETT A. McGUIRE, *Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA*; KELVIN LEE, *Accelerated Computing Systems and Graphics, Intel Corporation, Hillsboro, OR, USA*.

The interplay of the chemistry and physics that exists within astrochemically relevant sources can only be fully appreciated if we can gain a holistic understanding of their chemical inventories. Previous work by Lee et al. demonstrated the capabilities of simple regression models to reproduce the abundances of the chemical inventory of TMC-1, as well as provide predictions for the abundances of new candidate molecules. It remains to be seen, however, to what degree TMC-1 is a "unicorn" in astrochemistry, where the simplicity of its chemistry and physics readily facilitates characterization with simple machine learning models. Here we present an extension in chemical and physical complexity to an extensively studied hot star forming region, Orion-KL. Unlike TMC-1, the Orion-KL nebula is composed of several structurally distinct environments that differ chemically and kinematically, wherein abundances of molecules between components can have non-linear correlations that can cause the unexpected appearance or even the lack of unlikely species in various environments. A proof-of-concept study was performed to assess if similar regression models could accurately reproduce the abundances from the XCLASS chemical inventory obtained by the Herschel spectral survey. A new self-referencing embedded string (SELFIES) molecular embedder was adopted to account for vibrationally excited states and isotopologues. This additional complexity is considered with a hierarchical classification algorithm to indicate any relationships between environments with respect to the present species. Alongside the promising performance of our regression and classifier models, we attempted to fully capture the complexity of Orion-KL with increased efficiency using a neural network. The results of the classical models and neural network, as well as a discussion of their construction and performance, will be presented.