PLAGIARISM SCAN REPORT

Words	644	Date	January 21,2020
Characters	4322	Exclude Url	

0%	100%	0	32
Plagiarism	Unique	Plagiarized Sentences	Unique Sentences

Content Checked For Plagiarism

In the era of rapidly advancing frontiers of science and technology, the interaction of machine learning with various domains and fields is not at all surprising. Significant improvements and developments are being made not only in the mainstream research concerned with machine learning and artificial intelligence, but also in the fields of computational chemistry and molecular biology. In recent years, the advent of sophisticated deep learning and machine learning methods in molecular studies has grabbed the interest of various researchers and scientists. Their perspective, creativity garnished with a spoonful of knowledge and learning has paved its way to molecular machine learning. What is Molecular Machine learning? The profound use of machine learning and deep learning methods compiled with statistical analysis in chemical modeling, predicting the composition of the structures in molecular studies defines molecular machine learning. It defines a whole new approach to construct and reshape the properties of molecules. It's a buzz word in the current scenario, much more than one could ever describe. Overview Data is the primary resource for all the learning algorithms. Thus, molecular machine learning methods are capable enough in predicting the range of properties of a chemical structure based on the available data. The data for the molecule-based learning is highly heterogeneous and extremely expensive to gather in large amounts. Moreover, the imprecision and unstructured format requires enormous pre-processing to be performed along with a mastery in that domain. For example, the inorganic crystal structure database (ICSD) currently contains 188,000 entries, which have been checked for technical mistakes, but are still prone to human and measurement errors. The removal and identification of such algorithms is necessary for relevant conclusions. The pre-processing component has to be handled by digging out a way to represent this data in a form from which a machine could understand and learn. Molecules are the bunch of atoms chemically bonded together. Generally, the computer will not be able to understand it this way. Thus, the molecules are fed into the datasets in the form of SMILES (Simplified Molecular-Input Line-Entry System) strings and occasionally 3D coordinates are also used to represent the molecule features. SMILES have five syntax rules and predefined notations which can be incorporated using ASCII characters. The length sequence of ASCII characters for different molecules need to be the same as to ease out further computations. For this, the ASCII characters (particularly the chemical formula) are represented as a stream of vectors to the training or testing set, known as molecular featurization. Once the pre-processing part is performed well, there are many machine learning algorithms out there for distinct tasks. To name a few- Regression and Classification algorithms (Decision trees/ Naïve Bayes/ Random forest) and Reinforcement learning algorithms. The choice of the learner is important based on the domain knowledge, independent and target variables. Thus, by constructing a hypothesis, it can meet the demands in medical sciences. Applications Exploring the field of molecular studies using machine learning algorithms has many applications. The trending ones are listed below- • Mining and Predicting features/materials Machine learning has made it feasible to discover new drugs and the crystal structure of broader range of solids. The developments have led to remarkable improvements in medical sciences and research. Also, it has tremendous application in predicting the properties of large number of elements in order to examine their behaviour. • Increasing development time The slow progression in the field of medicinal science has caused many failures. On an average, it takes about 10 years to formulate a raw chemical structure to its final composition. But, accelerating their formation with the knowledge captured by the application of machine learning algorithms saves about a lot of time and money. Machine learning has outperformed the conventional molecular

computations. It has the potential to revolutionize the research in computational biology and molecular chemistry. Let's see, what the future beholds for it.

Sources	Similarity