NASA trains machine learning algorithm for Mars sample analysis

August 5 2024, by Matthew Kaufman

The Mars Organic Molecule Analyzer, aboard the ExoMars mission's Rosalind Franklin rover, will employ a machine learning algorithm to speed up specimen analysis. Credits: ESA

When a robotic rover lands on another world, scientists have a limited amount of time to collect data from the troves of explorable material, because of short mission durations and the length of time to complete complex experiments.

That's why researchers at NASA's Goddard Space Flight Center in Greenbelt, Maryland, are investigating the use of machine learning to

assist in the rapid analysis of data from rover samples and help scientists back on Earth strategize the most efficient use of a rover's time on a planet.

"This machine learning algorithm can help us by quickly filtering the data and pointing out which data are likely to be the most interesting or important for us to examine," said Xiang "Shawn" Li, a mass spectrometry scientist in the Planetary Environments lab at NASA Goddard.

The algorithm will first be put to the test with data from Mars, by operating on an Earth-bound computer using data collected by the Mars Organic Molecule Analyzer (MOMA) instrument.

The analyzer is one of the main science instruments on the upcoming ExoMars mission Rosalind Franklin Rover, led by ESA (European Space Agency). The rover, which is scheduled to launch no earlier than 2028, seeks to determine if life ever existed on the Red Planet.

After Rosalind Franklin collects a sample and analyzes it with MOMA, data will be sent back to Earth, where scientists will use the findings to decide the best next course of action.

"For example, if we measure a sample that shows signs of large, complex organic compounds mixed into particular minerals, we may want to do more analysis on that sample, or even recommend that the rover collect another sample with its coring drill," Li said.

NASA data scientist Victoria Da Poian presents on the MOMA's machine learning algorithm at the Supercomputing 2023 conference in Denver, Colorado. Credit: NASA/Donovan Mathias

Algorithm may help identify chemical composition beneath surface of mars

In [artificial intelligence](#), machine learning is a way that computers learn from data—lots of data—to identify patterns and make decisions or draw conclusions.

This automated process can be powerful when the patterns might not be obvious to human researchers looking at the same data, which is typical for large, complex data sets such as those involved in imaging and spectral analysis.

In MOMA's case, researchers have been collecting laboratory data for more than a decade, according to Victoria Da Poian, a data scientist at NASA Goddard who co-leads development of the machine learning algorithm. The scientists train the algorithm by feeding it examples of substances that may be found on Mars and labeling what they are. The algorithm will then use the MOMA data as input and output predictions of the chemical composition of the studied sample, based on its training.

"The more we do to optimize the data analysis, the more information and time scientists will have to interpret the data," Da Poian said. "This way, we can react quickly to results and plan next steps as if we are there with the rover, much faster than we previously would have."

Drilling down for signs of past life

What makes the Rosalind Franklin rover unique—and what scientists hope will lead to new discoveries—is that it will be able to drill down about 6.6 feet (2 meters) into the surface of Mars. Previous rovers have only reached about 2.8 inches (7 centimeters) below the surface.

"Organic materials on Mars' surface are more likely to be destroyed by exposure to the radiation at the surface and cosmic rays that penetrate into the subsurface," said Li, "but two meters of depth should be enough to shield most organic matter. MOMA therefore has the potential to detect preserved ancient organics, which would be an important step in looking for past life."

Future explorations across the solar system could be more autonomous

Searching for signs of life, past or present, on worlds beyond Earth is a major effort for NASA and the greater scientific community. Li and Da Poian see potential for their algorithm as an asset for future exploration of tantalizing targets like Saturn's moons Titan and Enceladus, and Jupiter's moon Europa.

Li and Da Poian's long-term goal is to achieve even more powerful "science autonomy," where the mass spectrometer will analyze its own data and even help make operational decisions autonomously, dramatically increasing science and mission efficiency.

This will be crucial as space exploration missions target more distant planetary bodies. Science autonomy would help prioritize data collection and communication, ultimately achieving much more science than currently possible on such remote missions.

"The long-term dream is a highly autonomous mission," said Da Poian. "For now, MOMA's machine learning algorithm is a tool to help scientists on Earth more easily study these crucial data."

The MOMA project is led by the Max Planck Institute for Solar System Research (MPS) in Germany, with principal investigator Dr. Fred Goesmann. NASA Goddard developed and built the MOMA mass spectrometer subsystem, which will measure the molecular weights of chemical compounds in collected Martian samples.

Provided by NASA