Jeff Schneider
Associate Research Professor
Robotics Institute
School of Computer Science
August 31, 2007 (3:30pm)

Research Topic
Machine learning

Research Problem
Can machine learning techniques be used to make the drug development process more efficient?

Problem Statement
Given behavioral data gathered from animal trials on specific drug compounds, construct a learning agent that can identify the class of a tested drug.

Operational Definitions
Animal trials: Experiments conducted on rats in a laboratory where the rats are given a drug compound (see below) and then continuously monitored for a set period of time in order to record their reactions.
Behavioral data: Videos continuously recorded throughout the animal trials that capture all of the animals' externally visible behavior (e.g. sleeping, running, eating, scratching, and so on).
Drug compounds*:* Combinations of chemical elements that are being tested to see if they have an effect on a given disease or condition; for example, the drug Paxil is a drug compound that has been shown to have an effect on the brain's serotonin's receptors that is beneficial to people suffering from depression.
Drug classes: Pre-existing categories of drug compounds that describe the effect the drug compound has in the body. For example, Paxil is classified as an antidepressant, meaning that it has an effect that lessens symptoms of depression.

Problem Description
The development process for new drugs is both time consuming and expensive. Years are spent identifying potential drug compounds and testing these compounds in the laboratory. Because of this, methods that can make these processes more efficient are in high demand. Schneider's work involves the use of machine learning algorithms to identify the class of a tested drug using data from animal trials, a process that can help identify which compounds are actually targeting the disease or condition they are being tested for. When using only data from tests of known drugs, the drug class can be identified with 66.2% accuracy. By contrast, the drug class can be predicted by chance alone with 7.7% accuracy. It can identify the specific drug 34.6% of the time, compared to the 0.7% accuracy expected by chance.

Computer Science Perspective
Computer science techniques are well-suited to solve this kind of problem via active learning. A computer can analyze a large dataset of behaviors and predict drug classes much more quickly than a human can, saving both time and money. This research also helps computer scientists develop more effective machine learning techniques.

Description of Disciplines Involved
This research involves collaborations with biologists and researchers in the pharmaceutical industry. Because this research focuses on drugs that target psychological conditions (such as depression), neurobiologists, psychologists, and psychiatrists are also involved.

Actively Involved Disciplines
Computer science, pharmacology, neurobiology, psychology.

Operational Definitions
Actively Involved Discipline: Any discipline from which one or more researchers made a significant contribution to the research design and interpretation of the results. Typically, the resulting research would add to the actively involved discipline's body of knowledge in some way, thus benefiting the discipline as a whole.

References
Presenter's homepage:
http://www.cs.cmu.edu/~schneide/
Auton Lab:
http://www.autonlab.org/

By
lajones

Updated
lajones