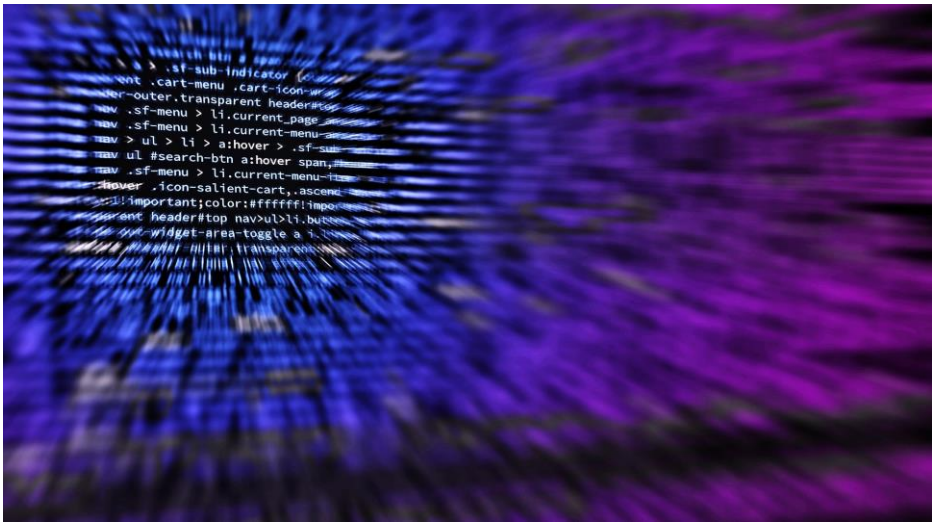
NovaUCD

Technology Licensing Opportunity

Data Cleaning for Predictive Machine Learning Systems

- Improving the prediction accuracy of artificial intelligence systems



Opportunity:

Artificial Intelligence and Machine Learning systems are currently trained on data that is entered by end-users either inputting data or by virtue of interacting with the intelligent system. This data is often very noisy, prone to error and needs to be pre-processed before it can be used. This "data cleaning" problem is now emerging as the real bottleneck in developing a trained artificial intelligence.

Researchers at University College Dublin and Teagasc have developed new algorithms and statistical techniques to clean noisy data sets which contain outlier data points prior to using such data sets for training a predictive machine learning system. These new techniques automate that data cleaning process, thus speeding up their implementation as well as improving the prediction accuracy of the artificially intelligent system.

Applications:

Artificially Intelligent Systems, Data Science, Big Data Analytics.

Key Features/Advantages:

- Structured, principled and automated methods to clean noisy data sets.
- Replaces manual and imprecise methods that are domain specific.
- Data cleaning is treated within a Bayesian statistical framework.
- Methods have been validated for a prediction models related to grass growth for precision agriculture.

FUNDERS: SFI Science Foundation Ireland, Insight, VistaMilk, Teagasc Agriculture and Food Development Authority

Value Proposition: Removes the data cleaning bottleneck and improves the accuracy of predictive machine learning models.

Market: Data Analytics, Artificial Intelligence, Precision Agriculture, Fintech.

Lead Inventor:
Professor Mark Keane
UCD School of Computer Science and Insight Centre for Data Analytics.

IP Status/Publication:
Non-patentable IP.

Contact:
Dr Tony Donnelly
Case Manager
Knowledge Transfer
t: + 353 1 716 3727
e: tony.donnelly@ucd.ie