

Bringing sequential feature explanations to life

Mokoena, Tshepiso
Lebogo, Ofentse
Dlaba, Asive
Marivate, Vukosi

ABSTRACT:

In many real world applications, human analysts are not only interested in the detected anomalies but are also interested in the reasons behind why they were flagged as anomalous. However, existing anomaly detectors provide the analysts with no information about what caused the anomalies. A sequential feature explanation(SFE) of a detected data point is an ordered sequence of features which are presented to the analysts, one at a time until the information contained in the set of already presented features is enough for the analysts to make a decision. However, SFEs are yet to be available on data analysis platforms that allow users to clean their data by filtering out anomalies. We present Quirk, the first of its kind user interactive anomaly detection system that adds the human analyst in the loop. Firstly, it loads the data from the user and then uses a user selected anomaly detection algorithm to identify the anomalies in the dataset. Secondly, for each flagged data point of interest to the user, it generates its SFE and provides a sequential visual presentation of the dataset along the features specified in the SFE with the flagged data point distinctly highlighted. Lastly, Quirk allows the analyst to provide it with feedback by labelling whether the flagged data points are truly anomalous or not. With this new information, it then computes a predictive model that will automatically flag anomalous data for future analysis. We present all of Quirk's functionalities and how it can be applied in real-world data analysis scenarios by presenting a use case of the system.