GTA ASSIGNMENT 5

Research Paper Summary

\

Smart-ML: A System for Machine Learning Model Exploration using Pipeline Graph

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India.
June 25-27, 2021

Done by -
N V Rakesh Reddy – PES1UG20CS566
Adithya M – PES1UG20CS621

Introduction

Humans are inherently good at interpreting visualizations and have been using visualizations as a medium for explaining a phenomenon or process to others. Therefore, the output of most of the approaches towards conferring human interpretability to ML models has been usually a visualization.

These visualizations aim to explain ML model behaviour to human users by highlighting the most contributing features. Studying interactions between features is equally important to facilitate knowledge discovery and providing new insights about the underlying physical phenomena. Existing visualizations mainly focusing on feature importance.

Some of the ML models are inherently interpretable. A few of the examples of inherently interpretable models are linear models, decision trees, and logistic regression. However, the majority of the complex ML models are opaque and lack human interpretability. These models do not explain why a particular decision outcome has been predicted for a given instance. This is termed as a lack of human interpretability.

Thanks to the advantages associated with interpretable ML models, a renewed interest has been observed among the research community over the past few years. Humans are known to be good at interpreting visualizations. As a result, in almost every work, aiming to confer human interpretability to an ML model, the idea has been to use a graph or visualization to present model behaviour or its decision-making process.

Proposed Methodology

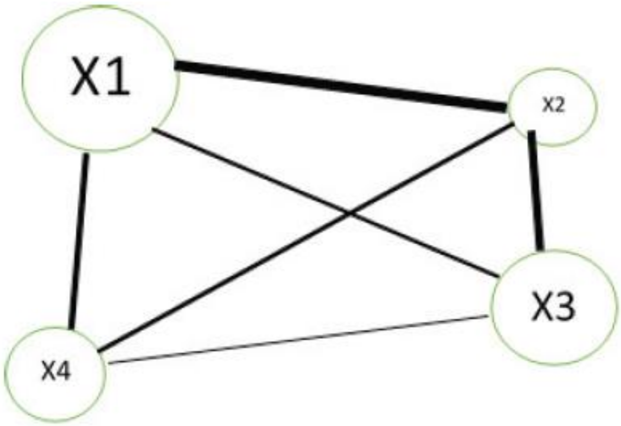
Most of the visualizations make use of a measure of feature importance. Feature importance is indeed an intuitive way of communicating behaviour of an ML model to its human users. However, studying the type and magnitude of the interaction between features help getting new insights about the underlying physical phenomena. It may lead to knowledge discovery especially in problem domains that are not well understood yet.

The aim is to create a visualization in the form of a graph that can represent feature importance as well as feature interactions in a single visualization. A human interpretable explanation to understand ML model behaviour in the form of a graph is an intuitive idea. It has an advantage that research and software tools available on network analysis can be used in interpreting ML models.

Feature - Importance and Feature-Interaction (FIFI) Graph:

- *Feature-importance* - The importance of a feature is measured as the increase in the prediction error after permuting feature values. A feature is termed as “important” if permuting its values increases the model error, else it is considered as “unimportant”.

- *Feature-interactions* - When features in a prediction model interact with each other, then the influence of the features on the prediction is not additive but more complex. The interaction between two features is the change in the prediction that occurs by varying the features, after having accounted for the individual feature effects.



A dummy FIFI Graph

Observations:

- Each node represents a feature
- Each edge represents the interaction corresponding to pair of features
- Size of the node is directly proportional to importance of the corresponding feature
- Width of an edge is proportional to the magnitude of the interaction between the corresponding pair of features.

Material And Methods Used:

1) Telco customer-churn (tcc) dataset:

- The ‘tcc’ dataset consists of customer records and is available publicly on Kaggle.
- The target variable was ‘Churn’ with values as ‘Yes’ or ‘No’.
- The dataset consisted of 7032 observations and 21variables.

Result: There were 1869 customers with churn status ‘Yes’ and 5163 customers with churn status ‘No’.

2) Freshmen(freshmen) Students dataset:

- The ‘freshmen’ dataset consisted of students enrolled at a private university in North India during the admission year 2018.
- The target variable was ‘JoiningStatus’ with values as ‘Joined’ or ‘Lost’.
- The dataset consisted of 13125 student records and 25 variables.

Result: There were 8374 students with Joining Status as ‘Joined’ and 4751 students with Joining Status as ‘Lost’.

3) ML algorithm:

For learning an ML-model, Random Forest (RF) algorithm was used. The RF algorithm was selected due to two reasons.

- It is a black-box algorithm and lacks human interpretability.
- Software tools are available to compute

4) Creating a FIFI graph

- To construct a FIFI graph for our RF model, the following steps were performed for features used in learning of the RF model:
  - Feature importance was computed for each feature
  - Feature interaction was computed for each pair of features
  - A visualization in the form of a graph was plotted where nodes represent feature and edges represent interactions
  - Node size was made proportionate to feature-importance
  - Edge width was made proportionate to feature-interaction
  - To improve interpretability, a sparse version of the graph was created by leaving out interactions with strength lesser than the mean interaction strength.

Results and Discussion:

Performance of RF model:
Baseline accuracy is prediction accuracy if the majority class is always predicted by an ML model.

Dataset	Baseline accuracy	Accuracy	AUC Value
tcc	0.734	0.791	0.822
freshmen	0.638	0.816	0.832

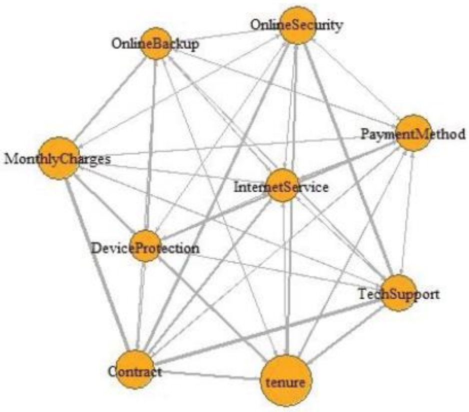
1) FIFI graph of RF model for 'tcc' dataset:

In the following table, the features are listed in descending order of importance. Each of these features is going to be a Node in FIFI graph, a unique ID was assigned to each feature.

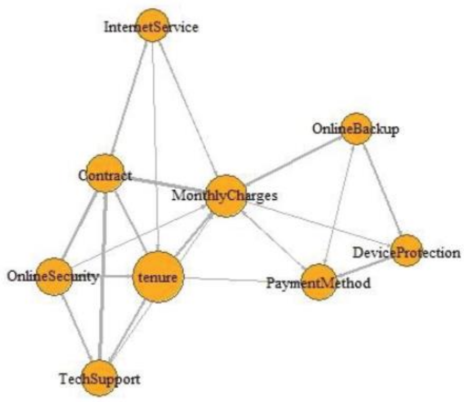
id	featureName	featureImportance
f01	tenure	1.9826
f02	MonthlyCharges	1.6380
f03	Contract	1.4963
f04	OnlineSecurity	1.4427
f05	TechSupport	1.4164
f06	PaymentMethod	1.3955
f07	OnlineBackup	1.2471
f08	InternetService	1.2864
f09	DeviceProtection	1.2447

The following table mentions the interaction strength of each feature with each of the other features except itself.

	f01	f02	f03	f04	f05	f06	f07	f08	f09
f01	0.0000	0.2674	0.3555	0.3999	0.3503	0.2972	0.1436	0.2027	0.0928
f02	0.4105	0.0000	0.4567	0.2115	0.2169	0.3134	0.3205	0.3076	0.2877
f03	0.2622	0.5093	0.0000	0.4320	0.5212	0.1787	0.0799	0.3476	0.0497
f04	0.3360	0.2506	0.4773	0.0000	0.3585	0.1671	0.1771	0.2230	0.0635
f05	0.2958	0.3099	0.4939	0.3414	0.0000	0.2048	0.1195	0.0915	0.0838
f06	0.1963	0.3121	0.1915	0.1927	0.2386	0.0000	0.2787	0.1902	0.2881
f07	0.1381	0.3892	0.0847	0.2102	0.1364	0.2553	0.0000	0.1537	0.3363
f08	0.3243	0.3030	0.2568	0.1710	0.0934	0.1885	0.1339	0.0000	0.0701
f09	0.0601	0.2144	0.0496	0.0534	0.0810	0.3413	0.2586	0.0843	0.0000



shows the FIFI graph for the RF model developed for 'tcc' dataset



shows a sparse variant of the FIFI graph.
It was obtained after deleting edges whose interaction strength is lesser than the mean interaction strength overall.

Observation:

It was observed that **'tenure'**, **'MonthlyCharges'**, and **'Contract'** are features contributing more in the decision-making process of the model.
The pairs (**'Contract'**,**'TechSupport'**), and (**'Contract'**, **'MonthlyCharges'**) were observed to have relatively strong interactions.

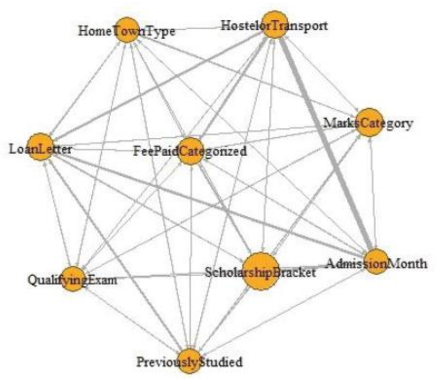
2) FIFI graph of RF model for ‘freshmen’ dataset:

id	featureName	featureImportance
f01	ScholarshipBracket	1.5608
f02	MarksCategory	1.1672
f03	HostelorTransport	1.1033
f04	FeePaidCategorized	1.1234
f05	LoanLetter	1.0949
f06	HomeTownType	1.0464
f07	PreviouslyStudied	1.0334
f08	QualifyingExam	1.0259
f09	AdmissionMonth	1.0145

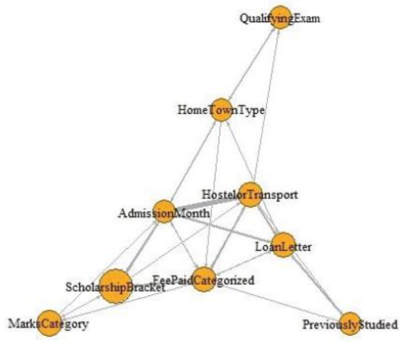
represents feature importance as per our RF model for ‘freshmen’ dataset.

	f01	f02	f03	f04	f05	f06	f07	f08	f09
f01	0.0000	0.2674	0.3555	0.3999	0.3503	0.2972	0.1436	0.2027	0.0928
f02	0.4105	0.0000	0.4567	0.2115	0.2169	0.3134	0.3205	0.3076	0.2877
f03	0.2622	0.5093	0.0000	0.4320	0.5212	0.1787	0.0799	0.3476	0.0497
f04	0.3360	0.2506	0.4773	0.0000	0.3585	0.1671	0.1771	0.2230	0.0635
f05	0.2958	0.3099	0.4939	0.3414	0.0000	0.2048	0.1195	0.0915	0.0838
f06	0.1963	0.3121	0.1915	0.1927	0.2386	0.0000	0.2787	0.1902	0.2881
f07	0.1381	0.3892	0.0847	0.2102	0.1364	0.2553	0.0000	0.1537	0.3363
f08	0.3243	0.3030	0.2568	0.1710	0.0934	0.1885	0.1339	0.0000	0.0701
f09	0.0601	0.2144	0.0496	0.0534	0.0810	0.3413	0.2586	0.0843	0.0000

Represents the interaction strength of each feature with each of the other features except itself.



shows the FIFI graph for the RF model developed for dataset



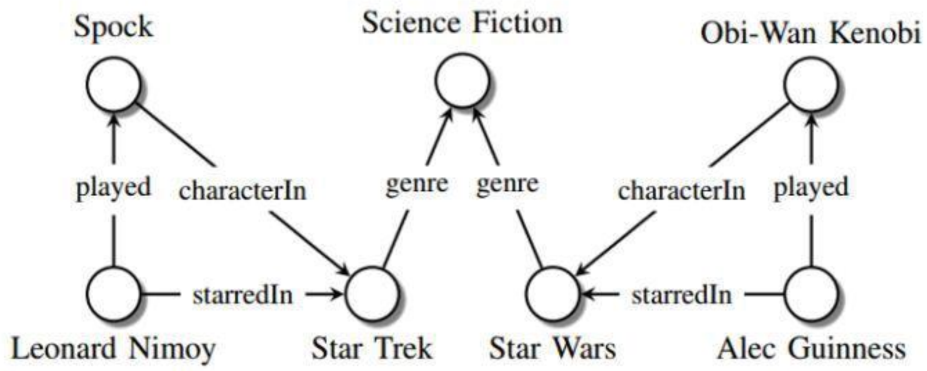
shows a sparse variant of the 'freshmen' FIFI graph.
The criteria used for the deletion of edges was the same as was used in 'tcc' dataset.

Observation:

- It was observed that **‘ScholarshipBracket’**, **‘MarksCategory’**, and **‘HostelorTransport’** are features contributing more in the decision-making process of the model.
- The pairs (**‘HostelorTransport’**, **‘LoanLetter’**), and (**‘HostelorTransport’**,**‘MarksCategory’**) were observed to have relatively strong interactions.

Comparison of FIFI Graph with a Knowledge Graph

- A knowledge graph is a way to organize and retrieve information in the form of a graph.
- It is used by search engines to improve their search results in response to user queries.
- Social networking sites and e-commerce sites are also using a knowledge graph to store and retrieve useful information. In a knowledge graph, nodes represent real-world entities and edges between nodes represent the existence of a relationship between nodes.
- Edge labels specify the type of relationship.



Conclusion

- Feature-importance and Feature-interactions can be plotted into a single visualization in the form of a FIFI graph.
    - It is a compact alternative to visualizing n+1 graph.

- Potential applications of a FIFI graph-based
    - It is useful in interpreting black-box ML models.
        - Stakeholders of an ML solution can interpret which features are affecting the decision-making process of the underlying ML model the most using this visualisation.
        - It is also useful in identifying prominent interactions between features.

    - It can be used in the design of interactive and interpretable ML systems where feedback from human users is part of the ML learning process.
        - Using an interactive version of a FIFI graph or a customized user interface, feedback from human experts can be taken in terms of feature importance as well as feature interaction by allowing the users to play with the size of a node and width of an edge.