Using Machine Learning to Inductively Learn Semantic Rules

Mohammed H. Jabardi\textsuperscript{1}, Asaad Sabah Hadi\textsuperscript{2}

\textsuperscript{1}College of Information Technology, the University of Babylon, Babylon, Iraq.

\textsuperscript{2}Information Technology Research and Development Center (ITRDC) The University of Kufa, Najaf, Iraq.

mohammed.naji@uokufa.edu.iq
asaadsabah@yahoo.com

Abstract. The Semantic Web and Machine Learning usually are seen as incompatible approaches toward Artificial Intelligence. A proposal presented for integrating the two paradigms and used data from Twitter regarding legitimate and fake accounts. Online Social Networks (OSN) such as Twitter have become a part of our lives due to their ability to connect people around the world, share documents, photos, and videos. OSN’s such as Facebook, Twitter and LinkedIn have approximately 500 million users over the world; this massive population of OSN causes different kinds of problems regarding data security and privacy. Unauthorised users infringe on the privacy of legitimate users and abuse names and credentials of victims by creating a fake account. We utilised Machine Learning to inductively learn the rules that distinguished a phoney account from a real one. We then implemented those rules in a Web Ontology Language (OWL) ontology using the Semantic Web Rule Language (SWRL). This integration provides the benefits of the data-driven ML approach combined with the explicit knowledge representation and the resulting ease of explanation and maintenance of the Semantic Web paradigm.

Keywords. Semantic Web, Machine Learning, Social Networks, Web Ontology Language, Semantic Web Rule Language, Decision Trees.

1. Introduction: The Technical Problem and Our Sample Problem Domain

1.1. Technical Approach: Integrating Machine Learning with Symbolic AI

The two most significant approaches to Artificial Intelligence (AI) at the current time are Symbolic AI and Machine Learning (ML). These two approaches are mostly opposites of each other. Symbolic AI began with Expert Systems. The goal of this approach is to represent knowledge in formats that are intuitive to human domain experts [1]. The Machine Learning approach, are based on algorithms that are sophisticated mathematical and logical models such as linear regression and Artificial Neural Networks. In the ML approach, the actual logic of the system not programmed by a developer but inferred from analysing large amounts of data called training sets.

The primary benefits of the symbolic AI approach are: [2]

- Knowledge represented in formats such as Rules and Ontologies that can apply to multiple problems across a domain.
• Systems are easy to maintain and extend because they defined in abstract formalisms (e.g., classes and rules) and intuitive to domain experts.
• Systems are transparent. Their reasoning is typically easy to understand, and they can provide explanation engines that provide English like explanations for decisions by analysing rule traces and axioms in an ontology.
• The primary benefits of the Machine Learning approach are [3]:
  • Problems such as pattern recognition and signal discrimination that could not be solved efficiently by the Symbolic AI approach are highly amenable to the ML approach.
  • Systems can take advantage of the massive amounts of data that organisations can collect via Online Social Networks (OSN) and eCommerce and can use this data to define efficient algorithms without requiring direct programming automatically. As more data becomes available, the algorithms achieve higher levels of accuracy.
  • The Symbolic AI approach typically does not handle inconsistencies in data well. The ML approach, on the other hand, is well suited to utilise large amounts of data. Even when such data has errors and outliers, the various ML algorithms will find the “best fit” to the data and minor errors in the data will have minimal effect on the resulting algorithm.
  • One of the most important goals that have been identified by many AI researchers is to integrate the two paradigms [4]. For example, Google scholar search found over 20,000 publications related to the interoperability of ML with symbolic technologies in 2017 [5]. An integrated approach could provide the ML ability to perform induction from large amounts of data with the symbolic AI potential for explanation and ease of maintenance. This paper describes an approach for integrating the two paradigms.

1.2. Problem Domain: Online Social Networking (OSN)
Recent reports indicate that networks like Facebook and Twitter infested with tens of millions of fake user profiles [6]. These counterfeit users can be individuals masquerading as someone else or automated computer bots designed to promote various political or social agendas such as supporting a particular political candidate or exciting a new movie with fake grassroots approval. These counterfeit users can collect personal details about real users and their friends and can endanger user safety both in the online and the real world. OSNs entice attention to malicious entities that attempt to exploit unintended vulnerabilities in OSNs. Threats to users include loss of privacy, identity theft, installation of malware, and Sybil attacks that impugn a real user’s reputation by falsely attributing opinions to them, and sexual harassment [7]. The data-driven ML approach combined with the explicit knowledge representation and the resulting ease of explanation and maintenance of the Semantic Web paradigm.

2. Related Works
Various approaches have been proposed in the integration of machine learning techniques and semantic web technologies. In[8] proposed flexible model by integrating Decision Tree with XML technologies for storing and retrieving knowledge. the model allows to add new ontological assertion to XML decision tree to gain for there reasoning. The metadata represented as RDF graphs and SPARQL query language used to retrieve it. [9] proposed SemTree which is a recommender system for specific domain knowledge(i.e choosing items from an overwhelmingly huge number of choices). SemTree model learned by integrating ontology with decision tree algorithm. SemTree model uses ontology and reasoner to generalize features of items semantically and enhancement the decision tree performance for building tree. SemTree is an accurate recommender system in domain knowledge. In [10] ontological engineering and decision tree classifier are used for generating new features from linked data. The rehearser showed that a decision tree classifier can be building by utilizing the hidden knowledge based on semantic web contains in ontology. In[11] presented new framework for the induction of logical decision trees by expanded the testing in each node using Description Logic concepts which permit to more division in the tree. The Web Ontology Language (OWL)used to express knowledge in the semantic web. In [12] the authors investigated interpretable model that increases the confidence of predicate values integrating
expert knowledge to model itself. The proposed model relying on an ontology to represent specific knowledge and on decision trees as a data-driven rule learning method. In [13] presented the Semantic Decision Tree Algorithm which used to mine the invisible knowledge in the semantic web data. Mining hidden knowledge from huge ontology becomes a necessary requirement due to increasing the data of semantic web in Linking data and search engines.

In [14] presented an algorithm to further improve the quality of ontology concepts with inductively derived decision trees. The concepts are symbolized by 2D regions and then describe how these regions can be used to match concepts of different but similar ontologies with each other. The prediction of the class-membership problem can be done by inducing either Evidential Terminological Decision Trees or Terminological Decision Trees, using Inductive Logic Programming (ILP) for solving the concept learning problem (i.e., concept for individuals of Semantic Web knowledge bases can be cast as a concept learning problem) [15]. In [16] presented OntoILPER which is a new system that integrating the ontology and Inductive Logic Programming (ILP) for acquisition entity and relation instances from textual data. The Named Entity Recognition (NER) and Relation Extraction (IE) both are based on machine learning technique as statical classification model. A recent review of the literature on this topic has been done in [17], predictive model has created for managing newly received orders in the manufacturing network. The model is based on ontological engineering and decision tree. WEKA application used for rule analysis and Semantic Web Rule language rules applied with the help of MATLAB programming.

3. Technologies

The two technologies we utilised for this work were Semantic Web technology and Machine Learning technology. Tim Berners-Lee initially defined the Semantic Web as one of the inventors of the Internet. The goal of Berners-Lee was always that the Internet would be indexed not by simple keywords as is currently the case but with semantics defined via ontologies that describe the sets, subsets, relations, and rules for various domains [18]. The current Internet was a necessary compromise based on the technology available at the time. Research primarily sponsored by the Defense Advanced Research Projects Agency (DARPA) in classification languages such as Loom [19]. Semantic Markup languages such as DAML [20], and knowledge-based technologies to represent large re-usable ontologies [21] led to the Web Ontology Language (OWL) [22] which is the W3C standard for defining ontologies and the Semantic Web Rule Language (SWRL) for defining rules [23].

Our training sets provided by over 1,000 examples of data on trusted and fake accounts collected from Twitter [24]. We have contacted the authors for providing us with a database.

3.1. Machine Learning

Machine Learning (ML) allows us to leverage the massive amount of data collected via sources such as OSNs and data warehouses. ML can use this data for prediction (e.g., how will a possible new data point align with models of the existing data) categorisation (e.g., discovering new categories in the data) and decisions informed by prediction and classification. ML is exceptionally well suited for applications where the problem is essentially one of signal detection in noisy data (e.g., face/eye recognition, recognising phonemes in a data stream of sound.) [25].

There are three types of ML approaches to learning: 1) supervised learning, 2) unsupervised learning and 3) reinforcement learning. In supervised learning we have the data (called features) that describe examples of the kind of input, then the system will receive and answers that correspond to each instance of the solution appropriate for that specific example. A supervised learning ML algorithm will use this data to develop a model that predicts future solutions for new sources of input data. The collection of features and solutions are called training sets.

On the other hand, with unsupervised learning, there are no solutions. Instead, the system discovers categories based on how data tends to be clustering together. The reinforcement learning is the third type of machine learning, unlike supervised learning and unsupervised learning; there are no pre-existing training sets. Instead, the system is continually giving
feedback based on its performance matching some desired goals such as an autonomous vehicle remaining on a path [26].

Problems in machine learning also can be divided into three categories based on the type of algorithm: Classification, Regression and Clustering. Classification problems are problems with categorical solution such as (“true” or “false”), (0 or 1) or (“yes” or “no”) which mean they belong to a particular set. Regression problems are problems wherein continuous values are inferred to match a mathematical model for a line or curve. Future costs can be predicted based on input variables matched to the numerical model. E.g., for the most straightforward kind of data, we can define a straight line in N dimensions and then predict a future value from N-1 input values matched to where that value would fit onto the hypothesised line.

Clustering problems can be found when the data needs to be organised to discover new patterns. One of the most common applications of clustering approach is to classify customers into clusters and then use those clusters to predict new products that the customer would find appealing based on the purchases of other customers in their same group. Many machine learning models deal with classification problems (the focus of this paper) such as Naïve Bayes, Logistic Regression, Random Forest and Decision Tree. Naïve Bayes and Logistic Regression are appropriate for more straightforward data when a line or curve can represent data. Whereas, if the data is very complex, the best choice is a Decision tree. The Random Forest technique is better for massive data sets [27]. In this paper, we used the Decision Tree approach.

3.2. Decision tree
In inductive machine learning, a decision tree acts as one of the most used methods. Decision Tree is forming by a set of training samples [28]. A Decision Tree is a tree-shaped diagram in the form of a ranching tree used to determine a type of action. The tree is organised to show how and why one of the options may be to the next opportunity, with the use of branching lines, indicating that each option is an option independent of the other choice. Leaf nodes indicate the final classification of a specific decision [29].

Decision Tree used to solve two kinds of problems, Classification (“true” or “false” or 0 or 1) and Regression (predicate values from a series of numbers or groups of data). A set of logical if-then conditions will determine to classify a problem, for instance, distinguishing between three types of animals depending on specific features. The Regression tree is using when the target variable is numerical or continuous [30].

The essential decision can be driven from other secondary choices based on other indicators. Secondary decisions can be made more specific depending on specific probability ratios. In this case, all possible alternatives to the decision are taken according to a certain probability. The primary decision, the secondary decisions and the associated sub-decisions are in their entirety the same as the tree and its branches. In this type of method, the tree is usually drawn according to its different directions, data and probability ratios, so that relations between branches and origin are also clarified. Thus, the decision tree is a quantitative and illustrative method of portrayal of the elements and relationships in which the problem is formed and under the various risk situations of nature. The tree diagram is the guide to indicate what section of the tree that can lead to the best results and less risk.

3.3. Semantic Web Technology
Rule-based Expert Systems primarily drove the first wave of the symbolic approach to AI. The most common example of the symbolic AI approach now is the Semantic Web.

The foundation for the Semantic Web is the Resource Description Framework (RDF). RDF provides a language for describing information as graphs of triples called assertions that consist
of a subject, property, and object. For example, WWW-Proposal hasAuthor Tim. The nodes in an RDF graph are the subject (in this example WWW-Proposal), and the object (Tim) and the edges are the properties (hasAuthor). The nodes and edges are all represented via Uniform Resource Identifiers (URI). A URI is a generalisation of the Uniform Resource Locator (URL) currently used to describe web addresses on the Internet. The difference is that URIs are not meant to be constrained to only Internet resources such as a web page but any resource including data that would typically be stored in a database and used by a program but not necessarily viewed in a web browser. The mapping from the current Internet to RDF is straightforward: web pages are RDF nodes and links are RDF edges [32].

RDF provides a powerful language to represent and reason about the Internet. However, the goal of the Semantic Web (where the term Semantic comes in) is to define higher levels of abstraction based on set-theoretic concepts such as sets, subsets, and relations. The language that provides rich semantic expressivity is the Web Ontology Language (OWL) [22]. OWL is built on top of RDF and adds concepts based on Description Logic (DL) to RDF. Description logic is a subset of First-Order Logic (FOL), and unlike FOL description logic is decidable. I.e., it is possible for a theorem prover to reason about an OWL model and to determine in a finite amount of time if the model is consistent.

The technology that led to OWL is KL-One and later languages such as Loom [19]. What came to be known as the KL-One family of languages used a theorem prover to take various axioms describing sets and relations (aka ?? classes and properties) and to use that information in two ways. First, to validate that the knowledge base (called an ontology in this paradigm) was consistent and had no contradictions, this is essential since in logic any formula can be derived from an inconsistency so once an ontology contains an inconsistency, it is useless for performing any automated inferencing. Second, the theorem prover (called a classifier in this type of language) would derive new inferences that can be deduced by the laws of logic from the user declared axioms.

With the vision of a Semantic Web, there was a new opportunity to utilise classification technology. Ontologies could describe vocabularies in common domains and bring structure and interoperability to web sites that far exceeded the keyword-driven websites that are still mostly the norm today, so the work facilitated by The Defense Advanced Research Projects Agency (DARPA) knowledge sharing initiative [21].

While OWL provides rich representation for data, it does not offer a way to implement rules, which provided by the Semantic Web Rule Language (SWRL) [13]. SWRL is a rule language that works with OWL. An SWRL rule consists of a series of statements on the left-hand side that are class and property expressions. These statements are implicitly universally quantified so that all variables (identified by placing a “?” in front of the variable). In the expression will fire the rule for every object that satisfies the variable. The statements on the right-hand side of the rule are all implicitly existentially quantified so that SWRL will make them valid for each object that meets all the expressions on the left-hand side. For example, if Human and Mortal are classes in the OWL ontology the SWRL rule: Human(?x) -> Mortal(?x) will fire for each individual in the ontology that is an instance of the Human class and assert that it is also an instance of the Mortal class.

The tool we used to develop our ontology was the Protégé ontology editor from Stanford University [31]. Protégé provides a highly customizable GUI environment for the definition of OWL ontologies, including SWRL rules. Protégé also supports several additional Semantic Web languages and tools that integrated with Protégé by defining them as “plug-ins” that can be loading in addition to the basic editor.
4. Approach

The proposed system consists of three main stages as shown in Figure 1. The first stage is data preprocessing and feature extraction using a data mining technique that involves transforming raw data into a useful format. In the second stage, the decision tree technique, which is one of the inductive machine learning used to build the decision tree rules. In the third stage, Ontology engineering used for creating the ontology and knowledge represented for utilising decision tree rules by converted to rules-based reasoner of semantic web rule language which used for detecting the fake Twitter account.

Semantic Web technology is used to implement a Machine Learning algorithm for detecting Twitter fake accounts from over 2,380 accounts. We applied our experiment on a dataset of Twitter accounts that are collected by “the Fake project”. The dataset consists of 2389 accounts and each account has 11 features as shown in Table 1.

| No | Features        | Description                                              |
|----|-----------------|----------------------------------------------------------|
| 1  | Statuses Count  | The number of Tweets (including retweets) issued by the user |
| 2  | Followers Count | The number of followers this account currently has        |
| 3  | Friends Count   | The number of users this account is following             |
| 4  | Favorites Count | The number of user’s accounts that this account is following |
| 5  | Listed Count    | The number of public lists that this user is a member of. |
| 6  | Retweet Count   | The number of retweets included in account                |
| 7  | Rep Count       | The number of user reputation included in account         |
| 8  | Hashtag Count   | The number of hashtags included in account                |
| 9  | URL Count       | The number of URLs included in account                    |
| 11 | Ment Count      | The number of user mentions included in account           |

Weka Gnu machine learning suite utilized to normalize the data and to apply induction for defining decision tree rules [32]. An example of supervised learning where a software engineer guides the algorithm through successive iterations, developing an algorithm that continually does a better job of correctly classifying an object (in this case classifying something as a Fake Account).

The Decision Tree algorithm in Weka application was applied and final Weka user interface shown in Figure 2, where the system has correctly classified 98% of the data.
The graphic for the resulting decision tree shown in Figure 3.

Stanford Protégé ontology editor was utilized to develop an OWL ontology for representing the classes and properties of the model. The main classes were Account and Agent. An Agent is any object that can own an Account. The most distinct subclass is Person, but there are other possible subclasses as well, such as Bots, Organizations, etc. Properties primarily defined in the class Account for describing attributes such as the number of followers, the number of friends, and other characteristics.
Figures 4 and 5 show views of the ontology as visualised in the WebVOWL ontology visualisation tool [33]. In WebVOWL dotted lines show subclass relations (with the arrow pointing to the superclass), straight lines show object and data properties (with the arrow pointing to the range of the property and the line beginning at the domain). The double arrows indicate inverse properties, e.g., hasEmployer and isEmployerOf are inverse properties. The red dots are simply a UI feature showing classes that have pinned for better visualisation.

We then developed an SWRL rule to represent the decision tree rules learned via induction by Weka. Due to the robust nature of SWRL, what were several rules for a decision tree could be implemented as a fewer SWRL rule.

5. Benefits
By integrating the results of our ML algorithm with Semantic Web ontology, we got many values of explicit knowledge representation and reasoning that come with Semantic Web ontologies.
There are many potential benefits of this integration. Since this is a small ontology developed primarily for this problem, we will only demonstrate a few of them. The real advantages are almost limitless and would be achieved by a more extensive ontology that was integrated with the Internet as well as with other Semantic Web systems.

5.1. Property Inference

Properties in OWL are necessary relations in First-Order Logic. As such, they provide all the standard descriptions and automatic inferencing that go along with relationships: transitivity, inverses, symmetry, identity, and many other relations.

For example, one object property in the ontology is hasAccount, which it refers to the relation between instances of the Agent class and instances of the Account class — for example, the Person user479 hasAccount ac479. The inverse property for hasAccount is the property isAccountFor. When this property defined in the ontology, the user does not need to determine its domain or range. All that is necessary is to declare that hasAccount and isAccountFor are inverses of each other. Since the domain and range for hasAccount are already known (respectively Agent and Account), the domain and range can be inferred by the reasoner to be respectively Account and Agent. Similarly, when the value for hasAccount defined for some Agent, the corresponding inverse value is deduced by the reasoner and vice versa. In this example, the fact that ac479 isAccountFor user479 automatically concluded with no additional work required by the user.

One example of how this could be used is with Description Logic queries. For example, the DL query:

\[ \text{Account and isAccountFor some (Agent and hasAccount some FakeAccount)} \]

This query would list all the Accounts that are owned by some Agent that has at least one FakeAccount even if the SWRL rule did not flag those accounts as fake accounts. These accounts would be dormant accounts that would be worth further scrutiny.

5.2. Explanations

It is also possible to use the explanation capability in Protégé to click on any instance of FakeAccount and generate a report as to why it classified as a FakeAccount. Figure 5 shows an example of a statement created for one such account. It shows the SWRL rule that was triggered and the property values that caused the rule to be triggered.

![Figure 6. Explanation generated for a Fake Account.](image)

6. Experimental Results and Discussion

In this paper, we described an approach where the rules for a decision tree to find fake accounts inferred from data collected on Twitter accounts. We then showed how this decision tree could be implemented in rules in the Semantic Web by using Ontology engineering and Semantic Web Rule Language (SWRL).

6.1. Result

In this section, we will illustrate the experimental results approach by using the data set (Fake Project IIT, CNR). Table.2 shows the standard factors metrics used in the evaluation process.
Table 2. Standard factors metrics.

| Factors             | Description                                                        |
|---------------------|--------------------------------------------------------------------|
| True positive (TP)  | The number of accounts correctly identified as Faked               |
| False positive (FP) | The number of accounts incorrectly identified as Faked             |
| True negative (TN)  | The number of accounts correctly identified as Trusted              |
| False negative (FN)| The number of accounts incorrectly identified as Trusted           |

To measure the quality of rules, we applying the following standard evaluation metrics [34]:

- **Precision**: the proportion of fake accounts those are fake, that is \( \frac{TP}{TP+FP} \)
- **Recall**: expresses relevant accounts that have been correctly detection, that is \( \frac{TP}{TP+FN} \)
- **F-Measure**: measures the quality of a prediction, that is \( 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \)
- **Accuracy**: state the number of accounts that are identified correctly in total = \( \frac{TP+TN}{TP+TN+FP+FN} \)

According to quality measures above, the result of the first experiment, using a decision tree and an ontology shown in Table 3 and Figure 7.

Table 3. Evaluation metrics results

| Approach   | FPR    | FNR    | Precision | Recall  | F-Measure | Accuracy |
|------------|--------|--------|-----------|---------|-----------|----------|
| Decision Tree | 0.040336 | 0.039933 | 0.989320 | 0.986622 | 0.988274 | 0.982419 |
| Ontology   | 0.040885 | 0.040880 | 0.986710 | 0.986710 | 0.986710 | 0.981582 |

Figure 7. Ontology statistics results.

6.2. Discussion

This paper aims to detect fake twitter accounts using the integration of machine learning (Decision Tree) and engineering ontology and SWRL rules. We use a dataset that published on MIB website. The decision tree is used as a classifier to detect twitter fake account from a real account. Figure 8 shows some decision tree rules used for classification.
Figure 8. Some of the decision tree rules.

We then developed an SWRL rule to represent the decision tree rules learned via induction by Weka. Due to the robust nature of SWRL what were several rules for a decision tree could be implemented as a fewer SWRL rule, as shown in Figure 9.

Finally, Decision Tree, correctly classified instances are 2347 account, and incorrectly cases are 42, while our Ontology correctly classified instances are 2345 and incorrectly classified examples are 48.

7. CONCLUSION AND FUTURE WORK
This study has gone some way towards enhancing our understanding of integration machine learning and Ontological engineering and rules-based reasoner of semantic web rule language. We also demonstrated some examples of how the knowledge representation and reasoning capabilities of the Web Ontology Language (OWL) could provide additional benefits in addition to the classification of fake accounts. There are many potential ways to advance this first step toward integrating the two paradigms. Some examples include:

- **Develop a richer ontology.** This ontology was reasonably fundamental. A richer ontology could consist of an ontology for security threats and rules regarding how to respond to them.

- **Extend to web-based objects.** One of the main benefits of the Semantic Web is that OWL ontologies can provide semantics for web pages. Semantic Web would be especially useful in this application where our entire domain is the domain of online social networking, so many of the objects in the ontology are web objects such as user pages and accounts.

- **Utilise ontologies that support fuzzy logic.** One of the significant advantages of machine learning is that it can help data with errors and inconsistencies. Ontology-based on logic, OWL does not handle inconsistent data or levels of confidence (e.g., saying an Account is a fake account with 80% confidence). There is research in progress to develop versions of OWL and other linked data graph models that do support logic with inconsistencies and levels of certainty, such as Fuzzy OWL [24]. These types of ontologies would provide additional benefits to
integration with machine learning algorithms and could support numeric algorithms such as regression which not based on decision trees.

- **Generalise this approach.** We have demonstrated a particular example of an approach that has applicability to many problems. This approach could be generalised and developed into software that automatically linked Weka and OWL, perhaps as a plug-in to the Stanford Protégé ontology editor.

**Acknowledgements**

This work was conducted using the Protégé resource, which is supported by grant GM10331601 from the National Institute of General Medical Sciences of the United States National Institutes of Health. We gratefully acknowledge the help provided by Michael DeBellis for feedback on the semantic web technologies.

**References**

[1] Hayes-Roth, F, Waterman, D, Lenat, D.: Building Expert Systems. Addison-Wesley, Reading Massachusetts (1983).

[2] Levesque, H, Brachman, R.: A Fundamental Tradeoff In Knowledge Representation and Reasoning. Readings in Knowledge Representation edited by Brachman and Levesque. Morgan Kaufmann. 1985.

[3] Ng, Andrew: What Artificial Intelligence Can and Can’t Do Right Now. Harvard Business Review. November 09, 2016.

[4] Lenat, D.: What AI Can Learn from Romeo & Juliet. Forbes: Cognitive World. July 3, 2019. https://www.forbes.com/sites/cognitiveworld/2019/07/03/what-ai-can-learn-from-romeo--juliet/#1d0031ad1bd0 Last Accessed: 7/8/2019.

[5] Kim, B, Finale, D-V.: Toward A Rigorous Science of Interpretable Machine Learning. Cornell University. arXiv:1702.08608. 2, Mar. 2017.

[6] Burns, J.: How Many Social Media Users Are Real People?. Gizmodo article. https://gizmodo.com/how-many-social-media-users-are-real-people-1826447042 2018/06/04. Last accessed: 2019/07/05.

[7] Varol, O. et. al. Online Human-Bot Interactions: Detection, Estimation, and Characterization. Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017).

[8] Diana Gorea and Sabin Buraga,"Integrating Decision Tree With XML Technologies",8th International Conference on Development and Application System, Suceava, Romania, May 25 – 27, 2006 A; Reif, G; Bernstein, Abraham; Gall, H (2008). SemTree: ontology-based decision tree algorithm for recommender systems. In: International Semantic Web Conference, Karlsruhe, Germany, 26 October 2008.

[9] Yordan Terziev,"Feature Generation using Ontologies during Induction of Decision Trees on Linked Data".CEUR Workshop Proceedings,ceur-ws.org,Vol-1733, paper-12

[10] J.L. Bal´czar et al."Induction of Concepts in Web Ontologies through Terminological Decision Trees", ECML PKDD 2010, Part I, LNAI 6321, pp. 442–457, Springer-Verlag Berlin Heidelberg

[11] Johnson I., Abé cassis J., Charnomordic B., Destercke S., Thomopoulos R. (2010) Making Ontology-Based Knowledge and Decision Trees Interact: An Approach to Enrich Knowledge and Increase Expert Confidence in Data-Driven Models. In: Bi Y., Williams MA.

[12] D. Jeon and W. Kim, "Development of semantic decision tree," The 3rd International Conference on Data Mining and Intelligent Information Technology Applications, Macao, 2011, pp. 28-34.

[13] Gajderowicz B., Sadeghian A., Soutchanski M. (2013) Ontology Enhancement through Inductive Decision Trees. In: Bobillo F. et al. (eds) Uncertainty Reasoning for the Semantic Web II. URSW 2010, URSW 2009, URSW 2008, UniDL 2010. Lecture Notes in Computer Science,
[14] Rizzo G., Fanizzi N., Lehmann J., Bühmann L. (2016) Integrating New Refinement Operators in Terminological Decision Trees Learning. In: Blomqvist E., Ciancarini P., Poggi F., Vitali F. (eds) Knowledge Engineering and Knowledge Management. EKAW 2016. Lecture Notes
[15] Zalan Mahmood Ayaz Khan, Salman Saeidlou & Mozafar Saadat (2019), Ontology-based decision tree model for prediction in a manufacturing network, Production & Manufacturing Research, 7:1, 335-349
[16] Zalan Mahmood Ayaz Khan, Salman Saeidlou & Mozafar Saadat (2019), Ontology-based decision tree model for prediction in a manufacturing network, Production & Manufacturing Research, 7:1, 335-349
[17] MacGregor, R.: Using a Description Classifier to Enhance Knowledge Representation. IEEE Expert 6 (3) June, 1991.
[18] Heflin, J, Hendler, J.: A Portrait of the Semantic Web in Action. IEEE Intelligent Systems 16 (2) July, 2001.
[19] Neches, R., et. al.: Enabling Technology for Knowledge Sharing. AI Magazine Volume 12 Number 3 (1991).
[20] Hebeler, J., Fisher, M., Blace, R., Perez-Lopez, A, Dean, M.: Semantic Web Programming. Wiley Publishing, Indianapolis, IN. 2009.
[21] W3C. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission 21 May 2004. https://www.w3.org/Submission/SWRL/, last accessed 2019/07/02.
[22] Ahmed El Azab, Amira M. Idrees, Mahmoud A. Mahmoud, Hesham Hefny,; ”Fake Account Detection in Twitter Based on Minimum Weighted Feature set”, International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:10, No:1, 2016.
[23] Richard O. Duda, Peter E . Hart, David G. Stork, A Wiley.: Pattern Classification (2nd Edition) Wiley-Interscience New York, NY, USA ©2000
[24] Stuart Russell,: Handbook of Perception and Cognition, Vol.14, Computer Science Division, University of California, Berkeley, CA 94720, USA
[25] Tom M. Mitchell: Machine Learning, McGraw-Hill Science/Engineering/Math; (March 1, 1997)
[26] M. R.; Tolun and S. M. Abu-Soud, “An Inductive Learning Algorithm for Production Rule Discovery,” Department of Computer Engineering Middle East Technical University. Ankara, Turkey, pp. 1–19, 2007.
[27] Savan Patel, Chapter 3: Decision Tree Classifier-Theory in Machine Learning, A journal for learning about Machine Learning, https://medium.com/machine-learning-101, last updated Jun 3, 2017
[28] Rastogi, R. & Shim, K.: A Decision Tree Classifier that Integrates Building and Pruning (2000) 4: 315. Kluwer Academic Publishers, Netherlands
[29] Protégé Web site. https://protege.stanford.edu/ Last accessed: 7/9/2019.
[30] Holmes, G, Donkin, A.: Weka: A Machine Learning Workbench. Proceedings of the Second Australia and New Zealand Conference on Intelligent Systems. Brisbane, Australia. 1994.
[31] Lohmann, S., Negru, S., Haag F., Ertl, T.: Visualizing Ontologies with VOWL. Semantic Web 7(4): 399-419 (2016)
[32] Bobillo F, Straccia. U.: Fuzzy Ontology Representation using OWL 2. International Journal of Approximate Reasoning, 2011
[33] Powers, David M W "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation". Journal of Machine Learning Technologies. (2011). 2 (1): 37–63.