ISSN: 2277-2685
IJESR/Sep.2020/Vol-10/Issue-3/1- 8
Mrs.K Vandana ***et.al.,*/International Journal of Engineering & Science Research**

Improving Modeling for Autonomous Machine Learning with a Task Ontology

Mrs.K Vandana[1], Mrs. Tasleem Sultana[2], Mrs. Matheen Sultana[3]

Assistant Professor[1,2,3]
Department of CSE
Global Institute of Engineering and Technology

Abstract:

Computing, big data, and machine learning algorithms are the building blocks of artificial intelligence, which has recently attracted a great deal of academic interest due to its potential to identify, learn, infer, and act upon external information across many domains. Nearly every sector is making use of AI technology right now, and a large number of machine learning specialists are trying to standardize and integrate different machine learning tools so that even those without a technical background may utilize them effectively. Ontology building for standardizing machine learning concepts and autonomous machine learning are additional areas of interest for the researchers. As part of this study, we outline a problem-solving process and categorize common processes in autonomous machine learning problem-solving as tasks. Our proposed approach to modeling autonomous machine learning makes use of a workflow-like procedure for executing machine learning tasks. The suggested machine learning model is based on task ontologies and proposes a strategy for grouping UML actions based on tasks. Furthermore, it will autonomously construct and augment machine learning models using transformation rules grounded on shared components and structures, including elemental linkages and processes.

I. Introduction

In recent times, artificial intelligence technology has emerged as a crucial tool for both academics and businesses. There are less obstacles to entrance into machine learning as most of the frameworks are open-source. Python, R, Scikit-learn, Caffe, TensorFlow, and Keras are some of the most used machine learning frameworks used in languages like R, Python, and Java. In this regard, a large number of ML specialists are attempting to integrate and standardize different tools in order to make them more accessible to fields where ML is not yet well-established. Autonomous machine learning, on the other hand, is in its early stages; some methods allow for the elimination of superfluous steps that are gradually improved upon in order to train the model and increase its precision. Using self-analysis, the autonomous machine learning tools provide the best algorithm for ML tasks and functions to find the hyper-parameter setting. Auto sklearn, Auto-Weka, H2o Driverless AI, and Google's Auto ML are some of the common technologies. This work presents the mechanism for autonomous machine learning, describes a typical problem-solving process for machine learning as tasks, and proposes a modeling way for employing task execution processes in machine learning. By defining a structure-based grouping technique of the UML (Unified Modeling Language) activities and implementing a function to automatically build models based on common components and structures, the autonomous machine learning modeling approach draws on the task ontology. To simulate autonomous machine learning, the suggested model makes use of pre-existing resources and generates new knowledge by relearning it.

Task ontologies

Ontology is defined in various fields depending on the field of applications. In the field of artificial intelligence, it is an explicit and formal specification of how objects and concepts described in the field of interest. In the Semantic Web, an ontology plays a very important role in processing, sharing, and reusing the knowledge for exchanging information between different databases. An ontology is also defined as an explicit description of concepts, attributes, constraints, and relationships between them on the domains. On the other hand, domain ontology can be defined as an 'explicit protocol for conceptualization' of the problem. A task ontology is defined as 'extracting and organizing the concepts and relations existing in the problem-solving process domain-independent'. In particular, a task ontology is a specification of the concept structure for the task execution process. Thus, the core concept is the subject of processing and the procedure of processing for a problem solving. In general, a person becomes a subject in a task ontology. However, in this paper, agents (programs) become subjects to perform the tasks. Expose ontology is an ontology for machine learning experiments. It is used in openML as a data structuring and data sharing (API) method. Machine Learning (ML) Schema is used to export all openML as linked open data. The DMOP ontology is explicitly designed to support data mining and machine learning. This covers the structure and parameters of predictive models, associated cost functions, and optimization strategies. Onto ontology provides a unique framework for data mining research.

Machine learning ontologies:

ML schema is a top-level ontology that provides classes, properties, and constraints for machine learning algorithms, datasets and experimentation suggested by the W3C (ML Schema community group). It can be easily extended and refined, and can be mapped to other domain ontologies developed in the field of machine learning and data mining. MEX vocabulary has been designed to solve the share of provenance information in a lightweight form. The extended PROV ontology provides a model for representing, capturing, and sharing provision information on the Web. This can enable the use of analytical data and code so that another person can reuse the results. The code and the markup language are written in a single file, and processed to create a document. A provenance meta information was proposed as a standard model of data management by W3C. The provenance information is also "information about entities, activities, and people involved in producing a piece of data which can be used to form assessments about its quality, reliability or trustworthiness". As a standard query language, SPARQL is a query language similar to SQL and stored in Resource Description Framework (RDF) for queries on data.

Autonomous machine learning modeling:

Autonomous machine learning modeling is the work for standardization and abstraction to the core of the components base on the meta information of the machine learning. The model consists of the task and process and saves as the method library (API). It defines into small units for modular and systematization of its components. The defined components redefine as a UML-based metamodel for the consistency, traceability, reusability, and implementation-ready between tasks and the results. So, the core class of the UML-based meta-model consist of tasks and processes. The Knowledge of the autonomous machine learning also describes a small task unit based on the MEXvocabulary. depicts a part of the knowledge of the object detection using the "YOLO" of the deep learning algorithm. The machine learning pipeline for object detection consists of data import, decision of attribute selection or schema, selection of learning model, construction of learning model, hyper-parameter setting, model training, measurement of model performance, and so on. In this way, the knowledge representation of the project unit is written as '.json' files using the mapping rules based on the machine learning schema and the vocabulary, and convert it into a UML-based meta-model.

II.LITERATURE SURVEY

"TensorFlow: A system for large-scale machine learning"

TensorFlow is a machine learning system that operates at large scale and in heterogeneous environments. TensorFlow uses dataflow graphs to represent computation, shared state, and the operations that mutate that state. It maps the nodes of a dataflow graph across many machines in a cluster, and within a machine across multiple computational devices, including multicore CPUs, general purpose GPUs, and custom-designed ASICs known as Tensor Processing Units (TPUs). This architecture gives flexibility to the application developer: whereas in previous "parameter server" designs the management of shared state is built into the system, TensorFlow enables developers to experiment with novel optimizations and training algorithms. TensorFlow supports a variety of applications, with a focus on training and inference on deep neural networks. Several Google services use TensorFlow in production, we have released it as an open-source project, and it has become widely used for machine learning research. In this paper, we describe the TensorFlow dataflow model and demonstrate the compelling performance that TensorFlow achieves for several real-world applications.

"Caffe: Convolutional Architecture for Fast Feature Embedding"
Caffe provides multimedia scientists and practitioners with a clean and modifiable framework for state-of-the-art deep learning algorithms and a collection of reference models. The framework is a BSD-licensed C++ library with Python and MATLAB bindings for training and deploying general purpose convolutional neural networks and other deep models efficiently on commodity architectures. Caffe fits industry and internet-scale media needs by CUDA GPU computation, processing over 40 million images a day on a single K40 or Titan GPU ($\approx$ 2.5 ms per image). By separating model representation from actual implementation, Caffe allows experimentation and seamless switching among platforms for ease of development and deployment from prototyping machines to cloud environments.

"Expose: An ontology for data mining experiments"
Research in machine learning and data mining can be speeded up tremendously by moving empirical research results out of people's heads and labs, onto the network and into tools that help us structure and filter the information. Expose, an ontology to describe machine learning experiments in a standardized fashion and support a collaborative approach to the analysis of learning algorithms. Using a common vocabulary, data mining experiments and details of the used algorithms and datasets can be shared between individual re-searchers, software agents, and the community at large. It enables open repositories that collect and organize experiments by many researchers. As can been learned from recent developments in other sciences, such a free exchange and reuse of experiments requires a clear representation.

"Task ontology: Ontology for building conceptual problem-solving models"
We have investigated the property of problem-solving knowledge and tried to design its ontology, that is, Task ontology. The main purpose of this paper is to illustrate a Conceptual Level Programming Environment (named CLEPE)as an implemented system based on Task ontology. CLEPE provides three major advantages as follows. (A) It provides human-friendly primitives in terms of which users can easily describe their own problem-solving process (descriptiveness, readability). (B) The systems with task ontology can simulate the problem-solving process at an abstract level in terms of conceptual level primitives (conceptual level operationality). (C) It provides ontology author with an environment for building task ontology so that he/she can build a consistent and useful ontology.

"Models for Representing Task Ontologies"
Knowledge is of general utility and should be captured thinking in reuse. A key idea underlining knowledge capturing for reuse is to consider that there are two major kinds of knowledge: domain and task knowledge. Ontologies can be used for representing both kinds of knowledge. while domain ontologies are broadly used and there are many proposals of models for representing them, the same does not occur for task ontologies. We propose the use of UML activity diagrams for capturing task control-flow, and UML class diagrams for capturing the knowledge roles involved in its activities. We also discuss the interrelationship between these two models and how they can be combined with domain ontologies in order to describe the knowledge involved in a class of applications.

III.SYSTEM ANALYSIS

1.1. EXISTING SYSTEM

The domain ontology can be defined as an 'explicit protocol for conceptualization' of the problem. A task ontology is defined as 'extracting and organizing the concepts and relations existing in the problem-solving process domain-independent'. In particular, a taskontology is a specification of the concept structure for the task execution process. Thus, the core concept is the subject of processing and the procedure of processing for a problem solving. In general, a person becomes a subject in a task ontology. However, in this project, agents (programs) become subjects to perform the tasks.

DISADVANTAGES OF EXISTING SYSTEM

- An autonomous machine learning is still in its infancy, and some techniques provide the ability to reduce the unnecessary tasks that are progressively refined to prepare the model and improve its accuracy.

1.2. PROPOSED SYSTEM

In the proposed system, we describe a typical problem-solving process for the machine learning as tasks, present their procedure, and propose the modeling method of an autonomous machine learning for using task execution processes. The modeling method of autonomous machine learning based on the task ontology define a structure-based grouping method of the UML (Unified Modeling Language) activities and implement a function to automatically generate models based on common elements and structures. The purpose of the proposed autonomous machine learning model is to model autonomous machine learning by reusing existing resources and producing new knowledge through relearning it.

ADVANTAGES OF PROPOSED SYSTEM

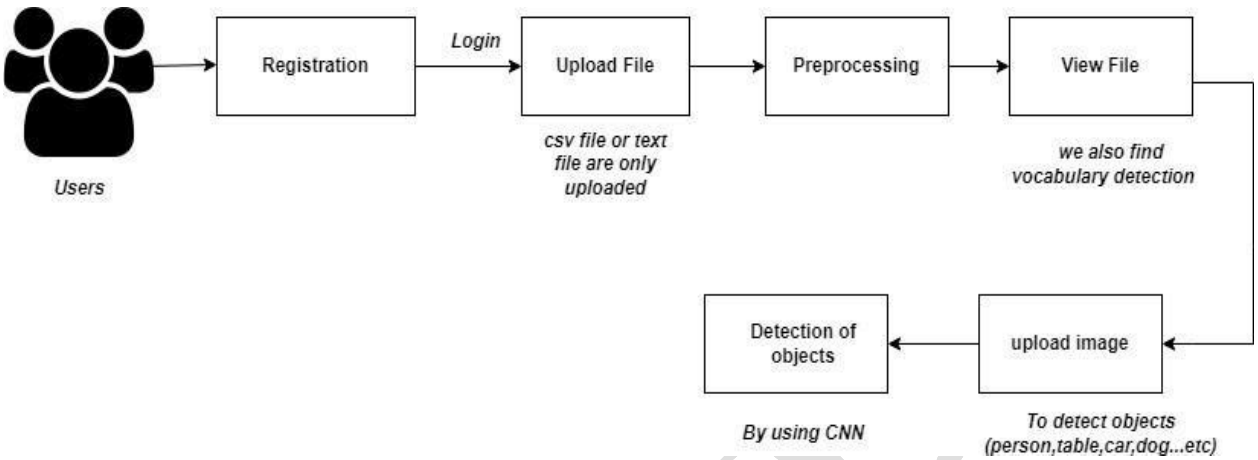
- ML schema is a top-level ontology that provides classes, properties, and constraints for machine learning algorithms, datasets and experimentation suggested by the W3C (ML Schema community group).
- It can be easily extended and refined, and can be mapped to other domain ontologies developed in the field of machine learning and data mining.

IV.SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture of autonomous machine learning modelling using a task ontology.

1.Foruserpage:



2.Foradminpage:

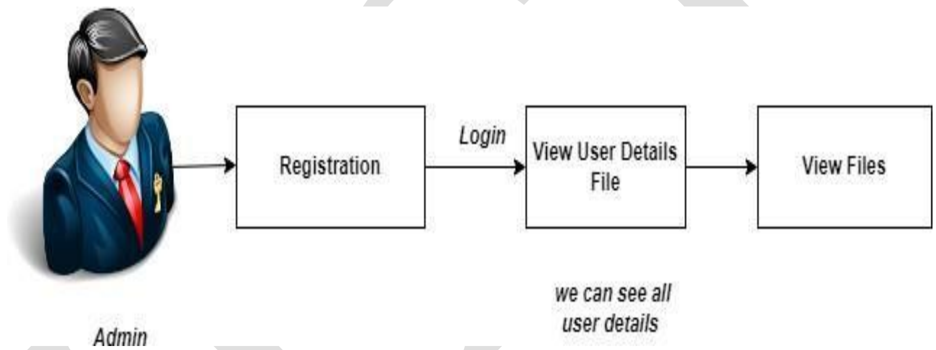


Fig.4.1.1SystemArchitecture

V. SYSTEM IMPLEMENTATION

5.1. MODULES

- USER
- ADMIN
- MACHINE LEARNING

User

User information and task descriptions for the entire experiment.we describe a typical problem-solving process for the machine learning as tasks, present their procedure, and propose the modeling method of an autonomous machine learning for using task execution processes. The collection of vocabulary is achieved by extracting words, a textbook or a machine learning library (API) tutorial, and then selecting keywords from the index and title of the textbook. The frequency of coincidence with the key word is calculated and labeled by a category item.

Admin

The aim of admin is to approve the machine learning users. the entire data must be gathered to admin. The design of autonomous machine learning model is presented based on the task ontology. The tools of autonomous machine learning provide an optimal algorithm for machine learning tasks and functions to determine the hyper-parameter setting through self-analysis.

Machine Learning

Machine learning refers to the computer's acquisition of a kind of ability to make predictive judgments and make the best decisions by analyzing and learning a large number of existing data. The representation algorithms include deep learning, artificial neural network, decision tree, enhancement algorithm and so on. The key way for computers to acquire artificial intelligence is machine learning. Nowadays, machine learning plays an important role in various fields of artificial intelligence. Whether in aspects of internet search, biometric identification, auto driving, Mars robot, or in American presidential election, military decision assistants and so on, basically, as long as there is a need for data analysis, machine learning can be used to play a role.

VI. RESULTS

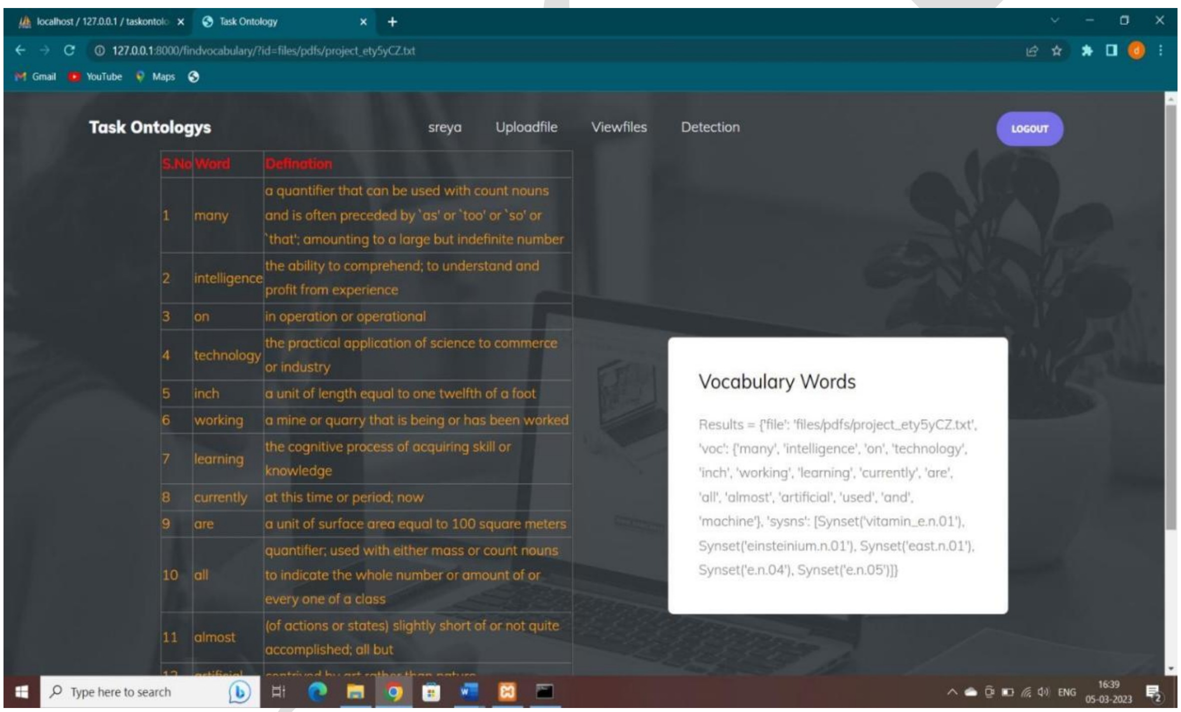


Fig.6.1VocabularyWords

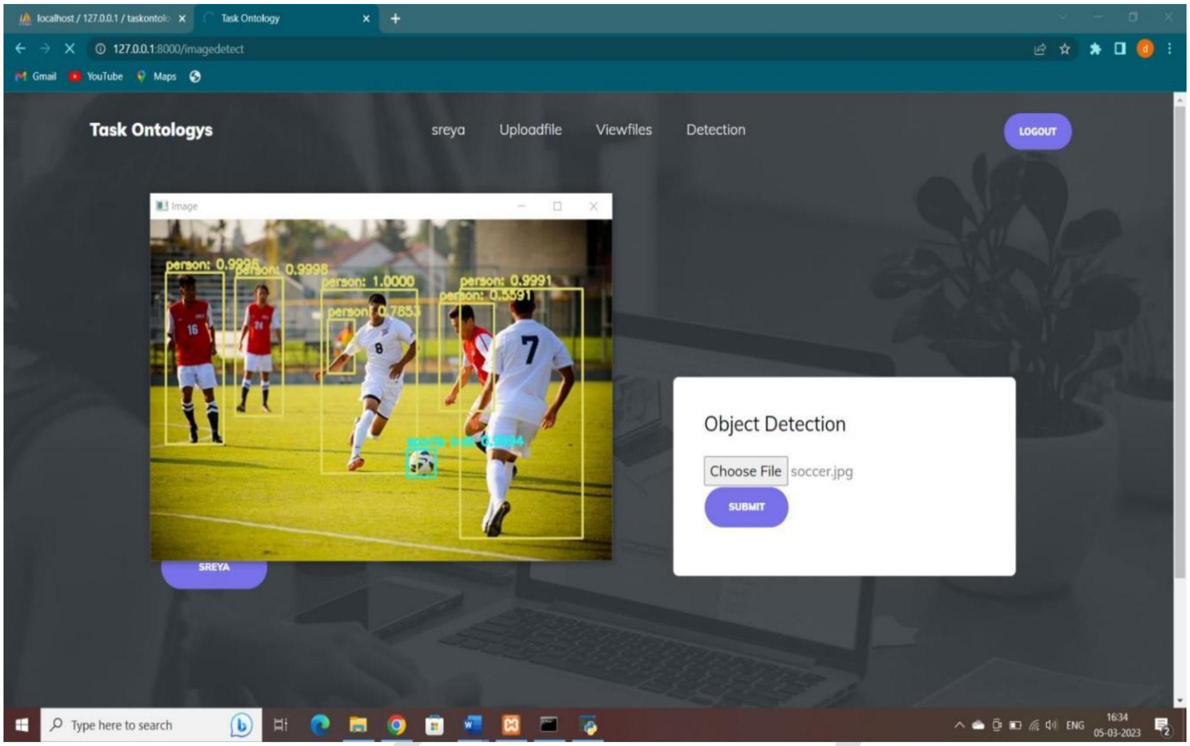


Fig.6.2 Image detection

VII.CONCLUSION AND FUTURE WORK

In this project, we extracted important keywords for constructing an ontology from papers and textbooks about machine learning. Moreover, we designed a task ontology based on the MEX vocabulary. We also studied workflow for the autonomous machine learning model. The proposed method is applicable for automatic workflow according to designated autonomous level. Therefore, the non-experts are capable of doing complex tasks using the proposed method and can easily implement the machine learning model in a specific application. Autonomous machine learning modeling using task ontology is a promising approach that can enhance the accuracy and efficiency of machine learning models. With the help of task ontology, machine learning models can better understand the context and requirements of the task at hand. This can lead to better model selection, hyperparameter tuning, and feature engineering, which can significantly improve the accuracy and generalization of machine learning models. In addition, autonomous machine learning modeling using task ontology can reduce the human involvement required in the modeling process, making it more efficient and scalable. This can save time and effort for data scientists, who can focus on more complex tasks. Future research can focus on improving the performance and scalability of task ontology-based machine learning models, as well as exploring new applications and domains where these models can be applied. In addition, research can focus on developing new techniques for representing and reasoning about tasks in machine learning, such as combining task ontology with other machine learning techniques such as reinforcement learning or transfer learning. In conclusion, autonomous machine learning modeling using task ontology has a promising future, and has the potential to revolutionize the way machine learning models are developed and applied in various domains.

REFERENCES :

[1] The paper "TensorFlow: A System for Large-Scale Machine Learning" was published in 2016 in the Operating Systems Design and Implementation journal and includes the works of authors A. Martín, B. Paul Barha, C. Jianmin, C. Zhifeng, D. Andy, D. Jeffrey, and others.

The information was retrieved from Keras's website on April 5, 2018.
the third"Caffe: Convolutional Architecture for Fast Feature Embedding," The Proceedings of the 22nd ACM International Conference on Multimedia, pp. 675-678, 2014, J. Yangqing, S. Evan, D. Jeff, K. Sergey, L. Jonathan, G. Ross, and others.

[4] In the 2011 Journal of Machine Learning Research, P. Fabian, V. Gaël, G. Alexandre, M. Vincent, T. Bertrand, G. Olivier, and others published "Scikit-learn: Machine Learning in Python," including pages 2825-2830.

5 April 2018 access to Theano from http://deeplearning.net/software/theano/.

"Autonomicity Levels and Requirements for Automated Machine Learning" [6] published by K. M. Lee, K. Y. Kim, and J. S. Yoo. Chapters 46–48 of the 2017 ACM International Conference on Research in Adaptive and Convergent Systems Proceedings.

Auto-Sklearn, available at https://automl.github.io/auto-Sklearn/stable/, version 7.

The Auto-WEKA program may be found at https://www.cs.ubc.ca/labs/beta/Projects/autoweka/.

AI for Hydrogen Vehicles,, Go to this URL: https://automl.github.io/stable/

[10] AutoML, a Google program, https://ai.googleblog.com/2017/11/automl-forlarge-scale-image.html, 2017.

I.M.: Intution Machine, 2017, C. Perez, The Deep Learning AI Playbook: A Strategy for Disruptive AI.

(Expose: An Ontology for Data Mining Experiments) by V. Joaquin and S. Larisa, presented at the Third Generation Data Mining Workshop at ECML PKDD 2010 [12].

"Task ontology: Ontology for building conceptual problem-solving models" (in the proceedings of the ECAI98 Workshop on Applications of Information and Knowledge Management), by I. Mitsuru, S. Kazuhisa, K. Osamu, and M. Riichiro