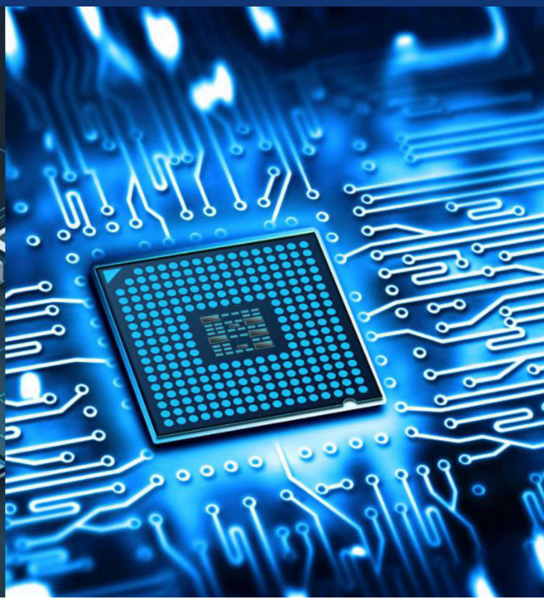
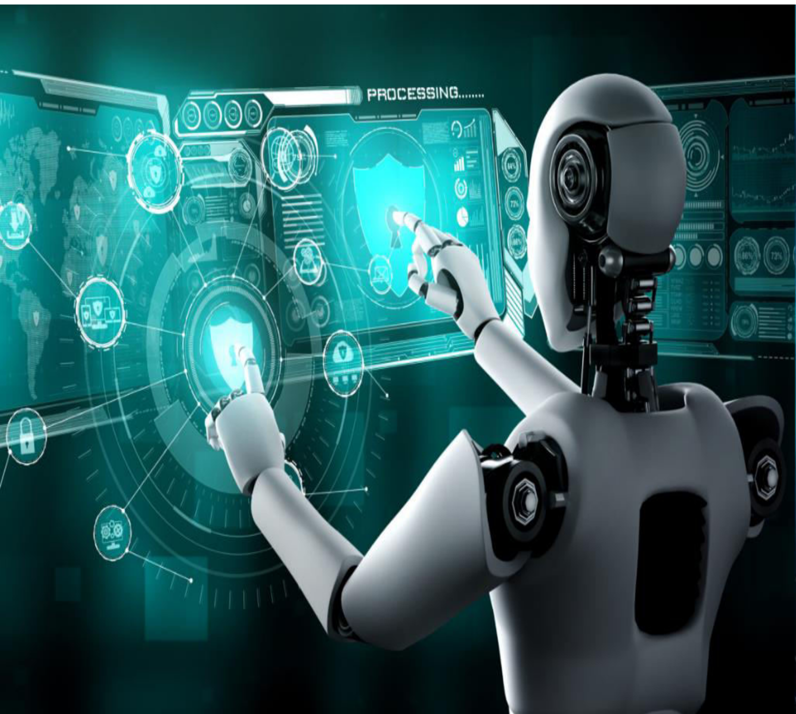
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 13, Issue 4, April 2025

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304170

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Task Ontology for Efficient ML Modelling

Sudhavali Adusumalli [1], P Sai Pavan [2], SK Nowshad [3], P Tejas Akshay [4], P Ravikiran [5]

Assistant Professor, Department of CSE, Eluru College of Engineering & Technology, Eluru, India[1]

B. Tech Student, Department of CSE, Eluru College of Engineering & Technology, Eluru, India[2,3,4,5]

ABSTRACT: Artificial intelligence technology that recognizes, learns, infers, and responds to external stimuli has recently attracted a lot of research interest. Information in a variety of domains by fusing big data, machine learning algorithms, and computing technologies. Nowadays, practically every industry uses artificial intelligence technology, and a large number of machine learning specialists are attempting to standardize and integrate different machine learning tools so that non-experts can use them with ease in their field. In order to standardize the concepts of machine learning, the researchers are also investigating autonomous machine learning and ontology construction. In this paper, we present a problem solving process and categorize common steps in autonomous machine learning problem solving as tasks. We suggest a way to model self-learning machines using a workflow of machine learning tasks. Our proposed machine learning model based on task ontology, sets up a way to group UML activities by task. It will also create and grow machine learning models on its own using rules to change common parts and structures (how elements connect and work together).

KEYWORDS: Machine Learning, Artificial Intelligence, TensorFlow and Ontology.

I. INTRODUCTION

Artificial intelligence technology has become one of the most essential tools in research and business context recently. Most of the machine learning frameworks are open-source, so the entry of barriers into machine learning are lowered. The typical machine learning frameworks include Tensorflow,Keras[, Caffe, Scikit-learn, and Theano implemented in programming languages such as Python, Java, and R. In this respect, many machine learning experts are working on integrating and standardizing various tools so that machine learning nonexperts can easily apply them to their domains. On the other hand, 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. 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. The typical tools include Auto sklearn, Auto-Weka, H2o Driverless AI and Google's Auto ML. In this paper, 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.

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. OntoMD 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. This model makes that objectives, optimizers, metrics, and layers in the Keras API are meta-model for deep learning.

Motivation: Our project aims to revolutionize machine learning with a Task Ontology Driven Autonomous Framework. This innovation will make machine learning more accessible and efficient. With autonomous learning, complex tasks will become easier to manage. Our framework will bridge the gap between experts and non-experts in machine learning. It will enable machines to learn and adapt on their own. This technology has vast potential across industries. We are excited to pioneer this cutting-edge framework. Our goal is to make machine learning more efficient, accurate, and accessible. By achieving this, we can create a brighter future for artificial intelligence.

Problem Definition: Current machine learning frameworks require extensive human intervention and expertise. This limits their accessibility and scalability. Manual feature engineering and hyperparameter tuning are time-consuming and prone to errors. Moreover, existing frameworks lack a standardized way of representing tasks and knowledge. This hinders the reuse and sharing of machine learning models. As a result, many industries struggle to adopt machine learning solutions. There is a need for an autonomous machine learning framework that can learn from data and adapt to new tasks. A task ontology-driven approach can provide a structured way of representing knowledge and tasks. This can enable machines to learn and reason autonomously. By addressing these challenges, our project aims to develop a revolutionary machine learning framework.

Objective of the work: The primary objective of this project is to design and develop a Task Ontology Driven Autonomous Machine Learning Framework. The framework aims to enable machines to learn and adapt autonomously from data. It will provide a structured way of representing knowledge and tasks using task ontologies. The framework will automate the machine learning process, reducing the need for human intervention. It will also enable the reuse and sharing of machine learning models across different tasks and domains. The project aims to improve the efficiency, accuracy, and scalability of machine learning solutions. It will also provide a user-friendly interface for non-experts to use machine learning. The framework will be evaluated on various datasets and use cases. By achieving these objectives, the project will contribute to the development of autonomous machine learning systems.

II. LITERATURE SURVEY

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, generalpurpose 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. A. Martín et al., described the TensorFlow dataflow model and demonstrate the compelling performance that TensorFlow achieves for several real-world applications.

J. Yangqing et al., presented a detailed note about Caffe . 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 generalpurpose 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 (≈ 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. Caffe is maintained and developed by the Berkeley Vision and Learning Center (BVLC) with the help of an active community of contributors on GitHub. It powers ongoing research projects, large-scale industrial applications, and startup prototypes in vision, speech, and multimedia.

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. **V.** Joaquin et al., presented 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. We therefore focus on the design of an ontology to express and share experiment meta-data with the world.

Mitsuru. I et al., 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. In this paper, firstly we briefly introduce the concept of task ontology. Secondly, CLEPE and its design principle is described. In CLEPE, one can represent his/her own problem solving knowledge and realize the conceptual-level execution.

III. SYSTEM DESIGN

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 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. The main with the existing system is, an autonomous machine learning is still in its infancy, and some techniques

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

provide the ability to reduce the unnecessary tasks that are progressively refined to prepare the model and improve its accuracy.

3.1 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.

3.2. MODULES:

The System Architeture mainly consists of three modules:
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.

The architecture diagram of the system is shown in figure 1.

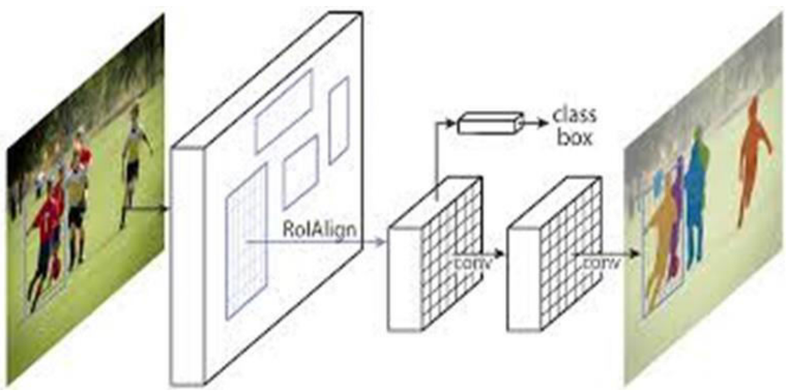


Fig 1: System architecture diagram

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. ALGORITHMS USED

4.1 DEEP LEARNING

Deep learning is a type of machine learning that uses artificial neural networks to learn from data. Artificial neural networks are inspired by the human brain, and they can be used to solve a wide variety of problems, including image recognition, natural language processing, and speech recognition.

Deep learning can be used in a wide variety of applications, including:

- Image recognition: To identify objects and features in images, such as people, animals, places, etc.
- Natural language processing: To help understand the meaning of text, such as in customer service chatbots and spam filters.
- Finance: To help analyze financial data and make predictions about market trends
- Text to image: Convert text into images, such as in the Google Translate app.

There are many different types of deep learning models. Some of the most common types include:

Convolutional neural networks (CNNs)

CNNs are used for image recognition and processing. They are particularly good at identifying objects in images, even when those objects are partially obscured or distorted.

Deep reinforcement learning

Deep reinforcement learning is used for robotics and game playing. It is a type of machine learning that allows an agent to learn how to behave in an environment by interacting with it and receiving rewards or punishments.

Recurrent neural networks (RNNs)

RNNs are used for natural language processing and speech recognition. They are particularly good at understanding the context of a sentence or phrase, and they can be used to generate text or translate languages.

4.2 NATURAL LANGUAGE PROCESSING

NLP, or Natural Language Processing, is a field of artificial intelligence focused on enabling computers to understand, interpret, and generate human language, allowing for more natural and intuitive interactions between humans and machines.

The primary goal of NLP is to equip computers with the ability to process and understand human language, enabling them to perform tasks like text analysis, translation, and even generate human-like text.

NLP combines computational linguistics (the rule-based modeling of language) with machine learning and deep learning techniques to analyze and understand the structure and meaning of text and speech.

4.3 YOLO & CONVOLUTIONAL NEURAL NETWORK

YOLO (You Only Look Once) is a real-time object detection algorithm that leverages a Convolutional Neural Network (CNN) to predict bounding boxes and class probabilities in a single pass, making it faster and more efficient than traditional CNN-based object detectors**.**

How it works:

- YOLO divides the input image into a grid and predicts bounding boxes and class probabilities for each grid cell.
- Unlike two-stage detectors (like R-CNN), YOLO performs object detection in a single stage, making it faster.
- YOLO treats object detection as a regression problem, predicting the bounding box coordinates and class probabilities directly.

Advantages:

Speed: YOLO is significantly faster than traditional CNN-based object detectors due to its single-stage architecture.

Real-time performance: YOLO is designed for real-time applications, making it suitable for tasks like self-driving cars and surveillance systems.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. RESULTS

The following figures present the sequence of screenshots of the results.

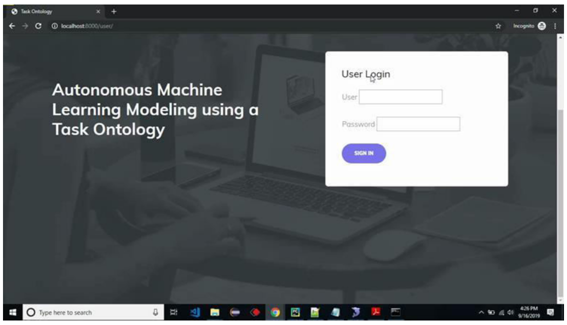


Fig 2: Login page

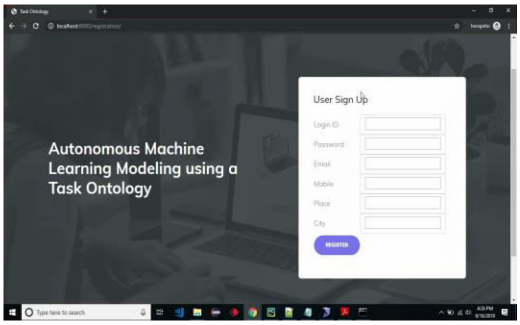


Fig 3: User register

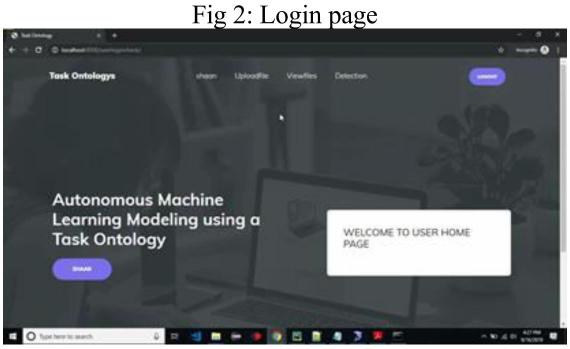


Fig 4: User home page

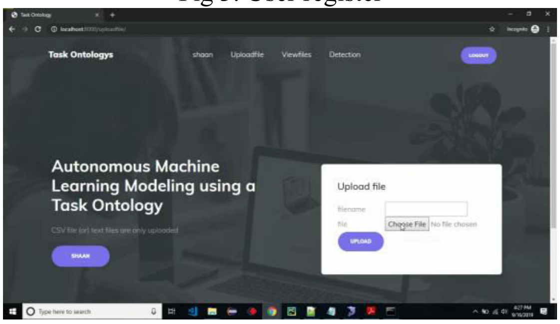


Fig 5: User upload file

Fig 6: User view file

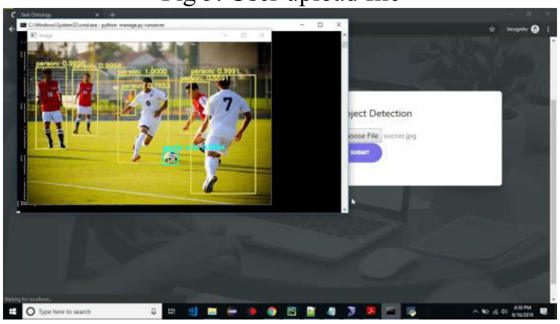


Fig 7: Detection of images

VI. CONCLUSIONS

In this paper, 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.

VII. FUTURE WORK

Future work will focus on extending the framework to support multi-task learning and transfer learning. We plan to integrate more advanced ontology reasoning techniques to improve the framework's decision-making capabilities. Additionally, we aim to develop a more user-friendly interface for non-experts to interact with the framework. We also plan to evaluate the framework on more complex datasets and real-world applications. Another direction for future work is to investigate the use of other knowledge representation formalisms, such as description logics. We also plan to explore the application of the framework in domains such as healthcare and finance. Furthermore, we aim to investigate the use of explainable AI techniques to provide insights into the framework's decision-making process. We also plan to develop a community-driven repository of task ontologies and machine learning models. By pursuing these directions, we aim to further advance the state-of-the-art in autonomous machine learning.

REFERENCES

[1] Martín, B. Paul Barha, C. Jianmin, C. Zhifeng, D. Andy, D. Jeffrey, et al., “TensorFlow: A System for Large-Scale Machine Learning," Operating Systems Design, and Implementation, Vol. 16, pp. 265-283, 2016.
[2] Keras, https://keras.io/accessed on April, 05, 2018.
[3] J. Yangqing, S. Evan, D. Jeff, K. Sergey, L. Jonathan, G. Ross, et al., “Caffe: Convolutional Architecture for Fast Feature Embedding,” Proceedings of the 22nd ACM International Conference on Multimedia, pp. 675-678, 2014.
[4] P. Fabian, V. Gaël, G. Alexandre, M. Vincent, T. Bertrand, G. Olivier, et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, pp. 2825-2830, 2011.
[5] Theano, http://deeplearning.net/software/theano/accessed on April, 05, 2018.
[6] K. M. Lee, K. Y. Kim, J. S. Yoo, “Autonomicity Levels and Requirements for Automated Machine Learning”. In Proceedings of the International Conference on Research in Adaptive and Convergent Systems, ACM. 2017, Sept.; 46-48.
[7] Auto-sklearn, https://automl.github.io/auto-sklearn/stable/
[8] Auto-WEKA, https://www.cs.ubc.ca/labs/beta/Projects/autoweka/
[9] H2o Driverless AI, , https://automl.github.io/ stable/
[10] Google's AutoML, https://ai.googleblog.com/2017/11/automl-forlarge-scale-image.html, 2017.
[11] C. Perez, The Deep Learning A.I. Playbook : Strategy for Disruptive Artificial Intelligence,I.M.:Intution Machine, 2017
[12] V. Joaquin, S. Larisa, “Expos´e: An Ontology for Data Mining Experiments”, Third Generation Data Mining Workshop at ECML PKDD 2010.
[13] I. Mitsuru, S. Kazuhisa, K. Osamu, M. Riichiro, “Task ontology: Ontology for building conceptual problem solving models”, In proceeding of ECAI98 Workshop on Applications of Ontologies and Problem-Solving model, pp.126-133, ECA, 1998.
[14] A. F. Martins, R. A. F. De, “Models for Representing Task Ontologies”, Proceeding of the 3rd Workshops on Ontologies and their Application, 2008.
[15] S. Kanjana, S. Maleerat, “Ontology Knowledge-Based Framework for Machine Learning Concept”, iiWAS '16 Proceedings of the 18th International Conference on Information Integration and Webbased Applications and Services, pp. 50-53, 2016.
[16] P. Gustavo Correa, E. Diego, L. Agnieszka, P. Panče, S. Larisa, S. Tommaso, V. Joaquin, Z. Hamid, “ML Schema: Exposing the Semantics of Machine Learning with Schemas and Ontologies”, ICML 2018 - RML Workshop.
[17] G. Eason, B. Noble, and I. N. Sneddon,"MEX Interfaces: Automating Machine Learning Metadata Generation,_ https://www.researchgate.net/publication/305143958, 2016. [18] E.Diego,N. M. Pablo, M.Diego, C. D. Julio,Z. Amrapali,L. Jens, “MEX Vocabulary: A Lightweight Interchange Format for Machine Learning Experiments”, SEMANTiCS 2016 Proceedings of the 12th International Conference on Semantic Systems, pp. 1724, 2016.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING