UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CMP 196 - ONTOLOGY ENGINEERING
MATHEUS VINÍCIUS TODESCATO

Modelagem do Domínio

Description of the Domain

Since graduation I have been working with research in the Machine Learning area, so I think it will be interesting to develop an ontology in this domain. Machine Learning is a sub-area of artificial intelligence that deals with an automatic process of searching for better representations of data. Unlike symbolic AI, which is based on rules, Machine Learning works with analytical engineering, performing training with data. This data may or may not be labeled, which can change the type of approach used for learning. Data analysis can be Statistical or Predictive which also changes the approach. Machine Learning can be divided into Unsupervised Learning (unlabeled data and statistical analysis), Supervised Learning (labeled data and predictive analysis) and Reinforcement Learning (works with feedback). For certain researchers, another category also exists, being Semi-Supervised Learning (or Active Learning) where we have labeled and unlabeled data together and a Supervised and with Reinforcement operation at the same time (that's why some researchers prefer to have this category within Reinforcement Learning). Algorithms need to be evaluated with metrics such as Recall, Precision, Accuracy. The main processes are training and testing. The training is for the model learn the rules of the data, while the test is for using these rules by predicting data that has not been seen before.

Modeling

UFO (initial modeling), GUFO (for protege) and OSA (Duarte, Bruno Borlini, et al. "Ontological foundations for software requirements with a focus on requirements at runtime." *Applied Ontology* 13.2 (2018): 73-105.)

Conceito	Supre identida de (O)	Carreg a identida de (I)	Rigidez (R)	Depend ência relacional (DR)	Dependência Existencial (DE1)	Meta-tipo
Machine Learning Algorithm	+	-	+	-	-	**Category**
Code	+	-	+	-	-	**Category**
TrainedModel					+	**Mode Universal**
Supervised Learning	+	+	+	-	-	**Kind**

Unsupervised Learning	+	+	+	-	-	**Kind**
Reinforcement Learning	+	+	+	-	-	**Kind**
Dataset	+	+	+	-	-	**Collective**
Data	+	+	+	-	-	**Kind**
Labeled Data	+	+	+	-	-	**SubKind**
Unlabeled Data	+	+	+	-	-	**SubKind**
Training					+	**Event**
Learning					+	**Event**
Testing					+	**Event**
Classification					+	**Disposition**
Regression					+	**Disposition**
Clustering					+	**Disposition**
Dimensionality Reduction					+	**Disposition**
Active Learning	+	+	+	-	-	**SubKind**
Passive Learning	+	+	+	-	-	**SubKind**
Algorithm	+	-	+	-	-	**Category**
Precision					+	**Quality**
Recall					+	**Quality**
Accuracy					+	**Quality**
Support Vector Machine **(all other types of algorithms)**		+	+	_	_	**Instance (Particular)**

Relações:

SVM(and all other types of algorithms) Instance-of Algorithm
Data Member-of Dataset
Labeled Data Subclass-of Data
Unlabeled Data Subclass-of Data
Dataset Member-of Machine Learning Algorithm
Active Learning Component-of Reinforcement Learning
Passive Learning Component-of Reinforcement Learning
Supervised Learning SubClass-of Machine Learning
Unsupervised Learning SubClass-of Machine Learning
Reinforcement Learning SubClass-of Machine Learning
Classification Type-of Supervised Learning

Regression Type-of Supervised Learning

Clustering Type-of Unsupervised Learning

Dimensionality Reduction Type-of Unsupervised Learning

Quality domains:

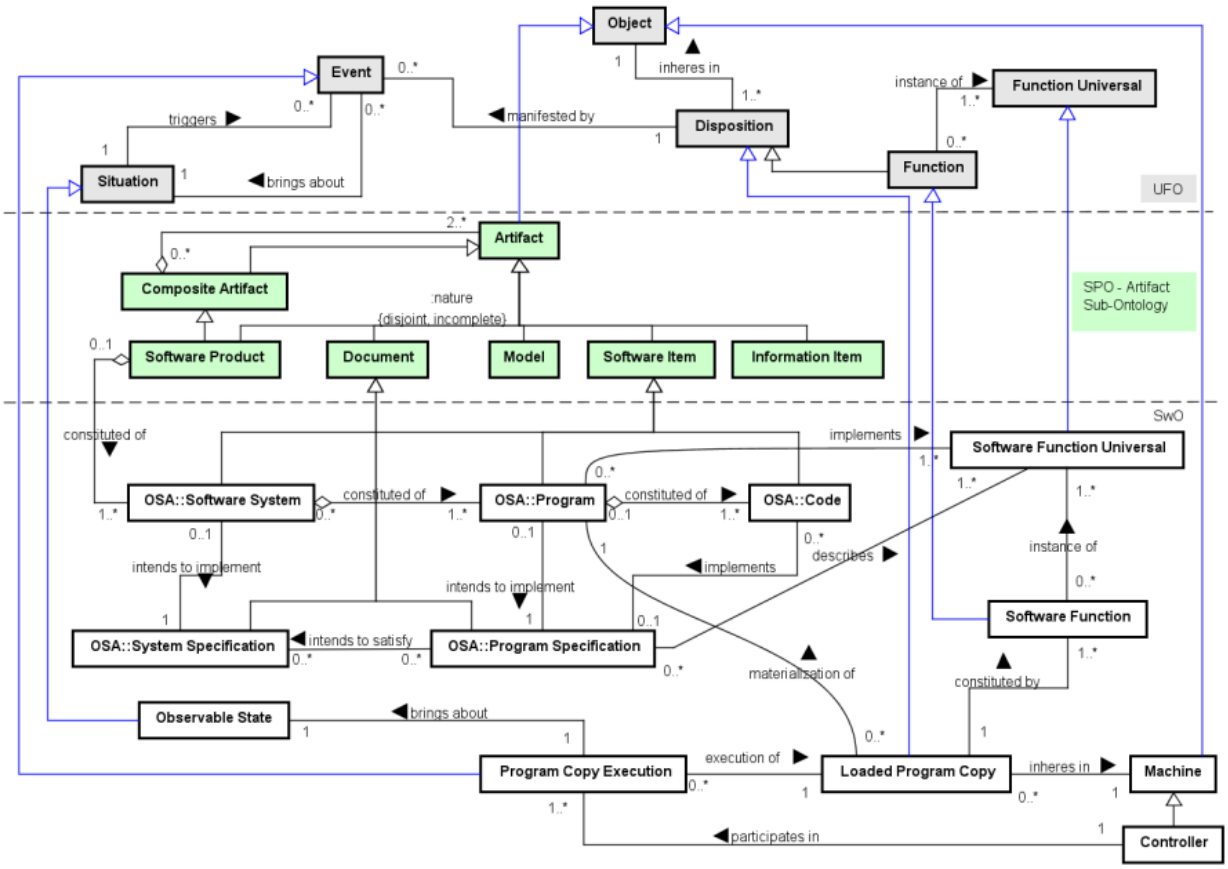
Precision: 0% a 100%

Recall: 0% a 100%

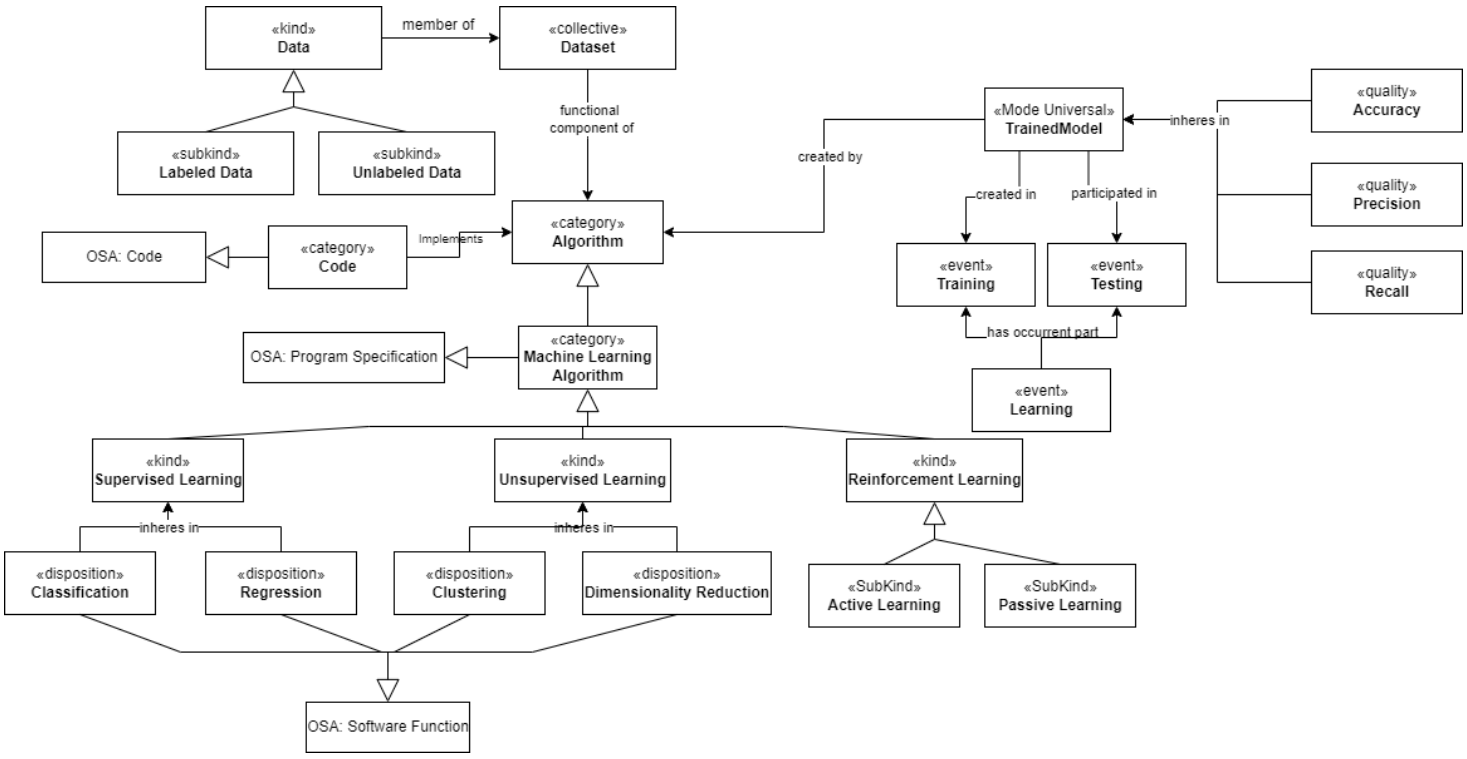
Accuracy: 0% a 100%

Conceitual - UML:

OSA:



ML_ontology:



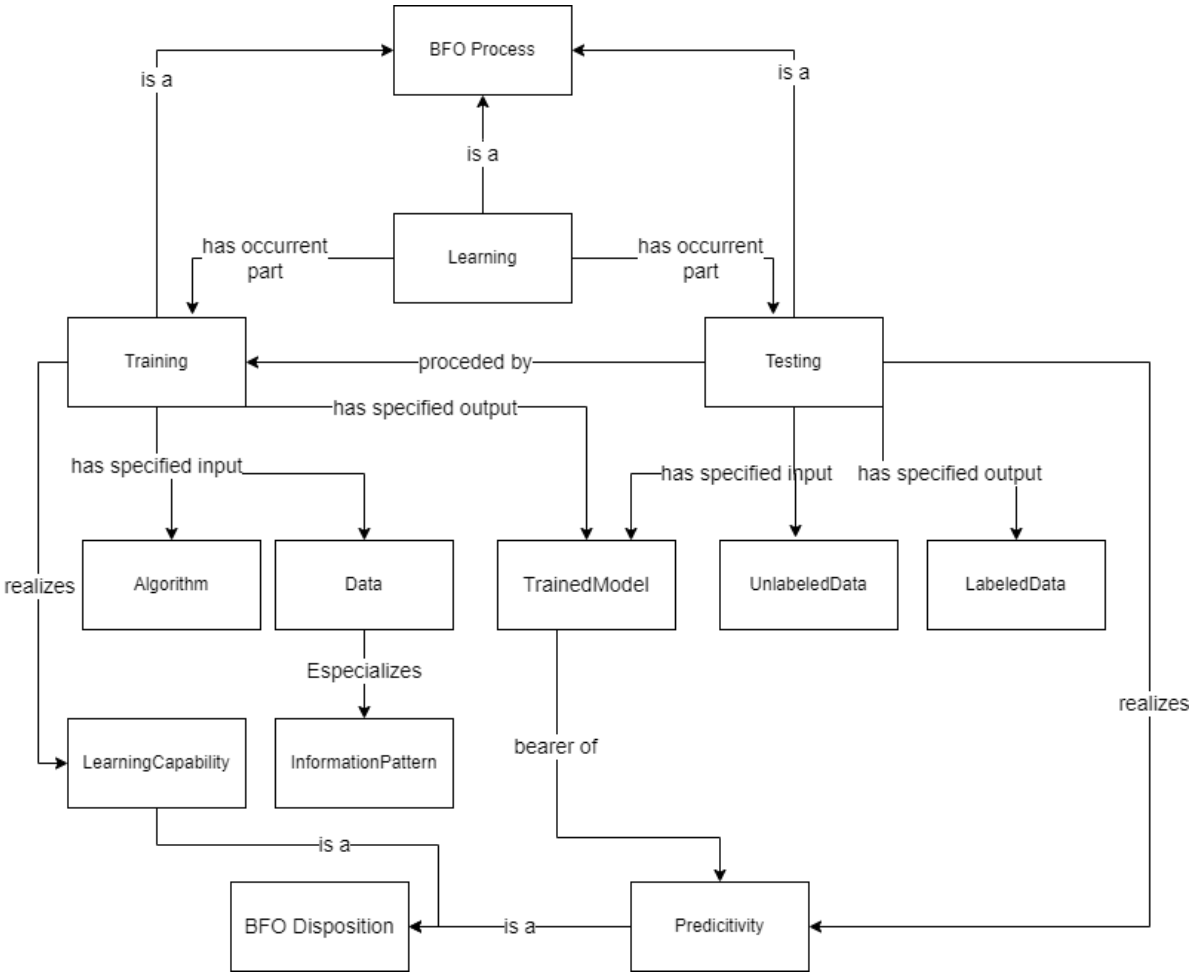
Ocorrent Modeling (BFO):

Learning

Occurrent Parts: Training, Testing

Participants of Training: Algorithm, Data.
Output of Training: TrainedModel.
Algorithm bearer-of LearningCapability
Data especializes InformationPattern

Participants of Testing: Trained Model, Unlabeled Data.
Output of Testing: Labeled Data.
Trained Model bearer-of Predictivity



Definition of Entities:

Conceito	Definição Semi-formal	Fonte	Exemplos	GUFO
OSA: Code	A sequence of machine instructions	Duarte, Bruno Borlini, et al. "Ontological foundations for software requirements with a focus on requirements at runtime." *Applied Ontology* 13.2 (2018): 73-105.	Code in Python	Object
OSA: Program Specification	A portion of Code can implement a Program Specification, which, in turn, may intend to satisfy a System Specification	Duarte, Bruno Borlini, et al. "Ontological foundations for software requirements with a focus on requirements at runtime." *Applied Ontology* 13.2 (2018): 73-105.	Algorithm	Object
OSA: Software Function	Directly related to what a system needs to do, i.e., to the functionalities it needs to provide for its users	Duarte, Bruno Borlini, et al. "Ontological foundations for software requirements with a focus on requirements at runtime." *Applied Ontology* 13.2 (2018): 73-105.	Classification of flowers	Object
Machine Learning Algorithm	Machine learning is a sub-field of computer science that aims to make computers learn	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.	SVM, Decision Tree…	Object
Supervised Learning	Supervised learning can be applied when the dataset contains a set of labels (possibly unitary) for every example.	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.		Object

Unsupervised Learning	When the dataset has no labels, that is, there is no previous classification of the examples, we apply an unsupervised learning algorithm. The goal is to infer classes or groups from the dataset without the help of the labels.	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.		Object
Reinforcement Learning	Reinforcement learning (RL) is an area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward. Reinforcement learning is one of three basic machine learning paradigms, alongside supervised learning and unsupervised learning. Reinforcement learning differs from supervised learning in not needing labelled input/output pairs be presented, and in not needing sub-optimal actions to be explicitly corrected. Instead the focus is on finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge).	Kaelbling, Leslie Pack, Michael L. Littman, and Andrew W. Moore. "Reinforcement learning: A survey." *Journal of artificial intelligence research* 4 (1996): 237-285.		Object
Dataset	Is a collection of data. Data sets can also consist of a collection of documents or files.	Snijders, Chris, Uwe Matzat, and Ulf-Dietrich Reips. "" Big Data": big gaps of knowledge in the field of internet science." International journal of	Dataset about COVID-19	FixedCollection

		internet science 7.1 (2012): 1-5.		
Data	Data are individual facts, statistics, or items of information, often numeric.	Stern, Roger, Ian Dale, and Sandro Leidi. "Glossary of statistical terms." *Google Scholar* (2000).	A tuple in a dataset	Object
Labeled Data	labels associated with every data	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.		Object
Unlabeled Data	Data without label	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.		Object
Active Learning	Active learning is a special case of machine learning in which a learning algorithm can interactively query a user (or some other information source) to label new data points with the desired outputs	Settles, Burr. "Active learning literature survey." (2009).		Object
Passive Learning	Learning to act from observational data without active environmental interaction	Ostrovski, Georg, Pablo Samuel Castro, and Will Dabney. "The Difficulty of Passive Learning in Deep Reinforcement Learning." *Advances in Neural Information Processing Systems* 34 (2021).		Object
Algorithm	Is a finite sequence of well-defined instructions, typically used to solve a class of specific problems or to perform a computation	Hill, Robin K. "What an algorithm is." *Philosophy & Technology* 29.1 (2016): 35-59.	Algorithm of binary search	Object

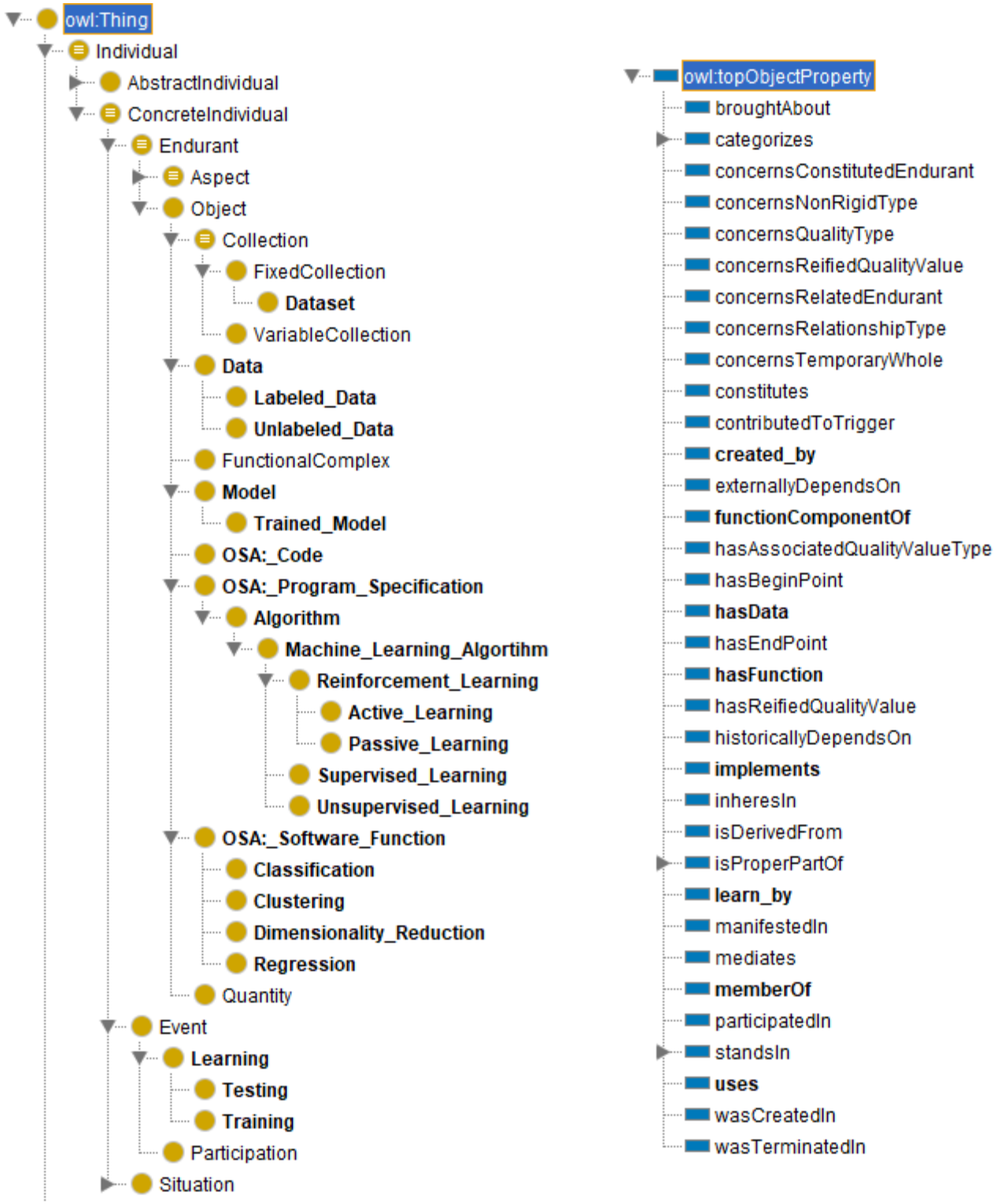
Classification	Classification is similar to a regression; the only difference is the labels that are discrete values	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.	Class of flowers prediction	Object
Clustering	Clustering is about discovering semantically related groups in an unlabeled dataset.	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.	Topics of news prediction	Object
Dimensionality Reduction	Dimensionality reduction plays an essential role in machine learning. Its goal is to decrease the number of features of a dataset.	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.	Features reduction in a dataset	Object
Regression	Regression is a type of supervised machine learning algorithm whose target variables are continuous values.	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.	Coin value prediction	Object
Learning	Learning is concerned with finding a right model, given a set of inputs and outputs.	Duarte, Denio, and Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.		Event
Training	Its the time to learner with the exemples and build a general model about this space that enables it to produce sufficiently accurate predictions in new cases.	Anzai, Yuichiro. *Pattern recognition and machine learning*. Elsevier, 2012.		Event
Testing	Time to evaluates the performance of the training model	Anzai, Yuichiro. *Pattern recognition and machine learning*. Elsevier, 2012.		Event
Trained Model	The model can be used	Duarte, Denio, and		Object

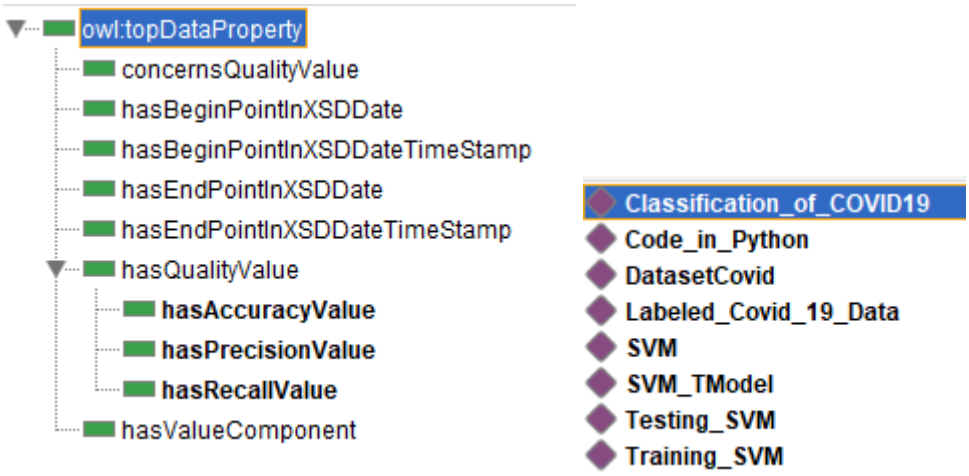
	to predict new outputs from new examples	Niclas Ståhl. "Machine learning: a concise overview." *Data Science in Practice* (2019): 27-58.		

Relações criadas:

Conceito	Definição Semi-formal	Exemplos
implements	A code implements an algorithm	Python code implementing SVM.
hasData	A Dataset has labeled or unlabeled data	Covid-19 dataset with labeled data
created_by	A trained model is created by an algorithm.	SVM trained model created by SVM algorithm
learn_by	An ML algorithm learns from data	Algorithm applied to Covid-19 dataset
uses	An algorithm type uses only a certain type of data.	Supervised Learning Algorithm Only Uses Labeled Data
functionComponentOf	Dataset is a functional Component of ML Algorithm	Dataset COVID-19 component of SVM algorithm
hasFunction	Function of the algorithm, of classifying, creating clusters…	Algorithm whose function is to classify data
memberOf	Data member of a dataset	A tuple from a dataset

Protege:


owl:Thing
Individual
AbstractIndividual
ConcreteIndividual
Endurant
Aspect
Object
Collection
FixedCollection
Dataset
VariableCollection
Data
Labeled_Data
Unlabeled_Data
FunctionalComplex
Model
Trained_Model
OSA:_Code
OSA:_Program_Specification
Algorithm
Machine_Learning_Algortihm
Reinforcement_Learning
Active_Learning
Passive_Learning
Supervised_Learning
Unsupervised_Learning
OSA:_Software_Function
Classification
Clustering
Dimensionality_Reduction
Regression
Quantity
Event
Learning
Testing
Training
Participation
Situation
owl:topObjectProperty
broughtAbout
categorizes
concernsConstitutedEndurant
concernsNonRigidType
concernsQualityType
concernsReifiedQualityValue
concernsRelatedEndurant
concernsRelationshipType
concernsTemporaryWhole
constitutes
contributedToTrigger
created_by
externallyDependsOn
functionComponentOf
hasAssociatedQualityValueType
hasBeginPoint
hasData
hasEndPoint
hasFunction
hasReifiedQualityValue
historicallyDependsOn
implements
inheresIn
isDerivedFrom
isProperPartOf
learn_by
manifestedIn
mediates
memberOf
participatedIn
standsIn
uses
wasCreatedIn
wasTerminatedIn



Nenhum problema relatado pelo Reasoner!