

Evolutionary Algorithm based automated design of Machine Learning pipeline

August 17, 2020

Aruna Viswanathan	CB.EN.U4CSE17610
Athisha R K	CB.EN.U4CSE17611
Manan Taneja	CB.EN.U4CSE17634
Jyothiswaroop R	CB.EN.U4CSE17653

Project Mentor: Dr. Shunmuga Velayutham C
Dept. of Computer Science and Engineering
Amrita School of Engineering
Amrita Vishwa Vidyapeetham

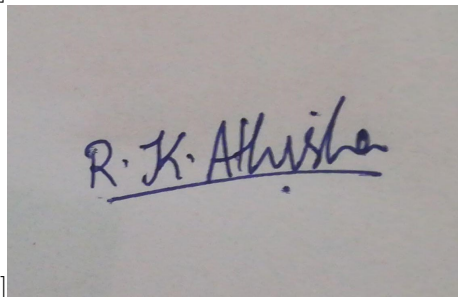
1 Abstract

Machine learning is a strategy for information investigation that automates expository model building. It is a branch of man-made reasoning in view of frameworks that can gain from information, recognize examples, and settle on choices with insignificant human mediation. Machine learning has made dramatic improvements in the past few years, but we are still very far from reaching human performance. Machine Learning is a difficult and time-consuming task that comprises of various steps including data preprocessing, feature extraction, choosing the right algorithm and tuning the hyperparameters, etc. The task becomes even more challenging since the designing of the specification is problem specific and requires domain-expertise. The Auto ML pipeline proposes an automated process of applying the machine learning technology to real-world business problems. In other words, it automates the entire pipeline right from obtaining the dataset (from diverse data sources) to the development of a customized machine learning model for the business. The Auto ML pipeline automates the selection of machine learning model by formulating it as a search and optimization problem. The main objective is that, the machine automatically searches through different models and combinations and selects the optimal model for the problem at hand. Evolutionary algorithms are stochastic search methods that mimic the metaphor of natural biological evolution. Evolutionary algorithms operate on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations

to a solution. At each generation, a new set of approximations is created by the process of selecting individuals according to their level of fitness in the problem domain and breeding them together using operators borrowed. Evolutionary Algorithms are robust optimizers which are well known for solving complex search problems. This project proposes to use Evolutionary Algorithm based automated design of Machine Learning pipeline for a given problem.

Arunaviswanthan

[a]



[b]



[c]



[d]

(a)Aruna Viswanathan (b)Athisha R K (c)Manan Taneja (d)Jyothiswaroop R