**SYSTEM ANALYSIS**

**EXISTING SYSTEM:**

In the existing system the data preprocess has dine with structured data. Even though data pre-processing consumes a large chunk of time in an ML pipeline, it is astonishing to see the inadequate amount of work done to automate it. For data preprocessing, it can be noted that while the data pre process approaches are adequate for structured data, work still needs to be done to assimilate on Structured data. We suggest the incorporation of data-mining methods as they can deal with such unformed data. This can allow AutoML pipelines to create models capable of learning from Internet sources. In feature engineering, it should be noted that most methods used until now adhere to supervised learning. However, dataset specificity is high, and therefore, AutoML pipelines should be as generic as possible to accommodate the diverse datasets. Therefore, a gradual paradigm shift towards unsupervised.

**DISADVANTAGES OF EXISTING SYSTEM:**

* Feature Generation is not up to the mark where domain experts excepted results.
* Most AutoML tools emphasize the performance but in the real world, that’s just one aspect being covered in machine learning projects. So the companies can’t compromise the computing plus storage specification sheet.
* CASH(Combined Algorithm Selection and Hyperparameter) problem considers model selection and hyperparameters optimization as a single hierarchical parameter.
* **Algorithm**: SmartML,J48,C50

**PROPOSED SYSTEM:**

The proposed System aims at providing an overview of the advances seen in the realm of AutoML in recent years. We focus on individual aspects of AutoML and summarize the improvements achieved in recent years. The motivation of proposed system stems from the unavailability of a compact study of the current state of AutoML. While we acknowledge the existence of other surveys, their motive is to either provide an in-depth understanding of a particular segment of AutoML, provide just an experimental comparison of various tools used or are fixated towards deep learning models.

**ADVANTAGES OF PROPOSED SYSTEM:**

* We segment the AutoML pipeline into parts and review the contributions in each of these segments.
* We explore the various state-of-the-art tools currently available for AutoML and evaluate them.
* We also incorporate the advancements seen in machine learning which seems to be overshadowed by deep learning in recent years.

**Algorithm**:H2O-AutoML, LinearRegression, Gradient Boosting Regressor