**Review of SIMON, a piece of open-source software**

References

Tomic, A., Tomic, I., Waldren, L., Khatri, P., Pollard, A, J., & Davis. M, M. (2021, January).

SIMON Open-source knowledge Discovery Platform.

Website\*: <https://doi.org/10.1016/j.patter.2020.100178>

Purpose of paper

The purpose of this work is to give an overview of SIMON (Sequential Iterative Modeling Over Night), a pattern identification and information extraction tool. The primary advantages of SIMON, the motives for its deployment, its features, and outcomes with diverse data sets, and how SIMON assists in the identification and interpretation of complicated biological data are all highlighted.

Research Design or Strategy

This paper's study strategy covers many major concerns, including the detection of complicated, quickly changing biological data, the extraction of complex data, and the analysis of such complex data using a single tool that requires no coding and may help anybody acquire more accurate findings. SIMON is one such piece of software (Sequential Iterative Modeling Over Night). Machine learning, a subfield of artificial intelligence, allows for the unbiased extraction of large amounts of diverse data. Even for experienced informaticians, incorporating machine learning demands a significant amount of effort; hence, software is developed to provide a consistent ML method for data pre-processing, data splitting, constructing predictive models, and producing good visualization results. "SIMON" (Sequential Iterative Modeling "Over Night") is an open source application developed by authors to address all of these challenges. It provides a user-friendly, standard interface for constructing and evaluating prediction models using a variety of machine learning approaches. There are presently 182 algorithms available in this application, which is available in two versions: single mode and server mode. "Through the creation of training, validation, and test sets, SIMON delivers a single approach for model training, hyperparameter tuning, and model assessment." Single mode supports Windows, Mac, and Linux and has the following main features and procedures: 1. Creating prediction models makes data selection easier. Data classification, including high-dimensional data sets, and model selection 3. SIMON is one of the most effective tools for identifying complex biological data. Comparable features include feature selection, exploratory analysis, cloud compatibility, the ability to distribute work evenly across all connected computers, parallel processing, and regular source code updates.

**Results, Discussion, Conclusion of paper**

**Results:** SIMON (Sequential Iterative Modeling "Over Night") program provides various advantages, according to the research and review cited above. It works with any operating system, requires no coding knowledge, splits work simply, and can be worked from the cloud. It can also collect high-quality heterogeneous data for use in knowledge extraction, pattern recognition, and other applications.

**Discussions:** SIMON (Sequential Iterative Modeling "Over Night") software is particularly powerful in data mining and offers pattern identification and data extraction features for high-quality heterogeneous data. Aside from managing data from the cloud in the SIMON server version, there are risks to data privacy and security because it is open source. Data pretreatment processes must be performed manually in this program, and because the C++ language is employed, there are risks to data privacy and security.

**Conclusions:** The major result of this article is how the program SIMON extracts and analyzes heterogeneous complicated biomedical data (Sequential Iterative Modeling "Over Night"). It also explains the software's features and how accurate it is at analyzing data and providing desired results fast and effortlessly.

Contribution from paper

# The authors' major goal in this study is to make complex heterogeneous complex data simple to analyze and to reduce time and money in data analysis by utilizing SIMON, an open-source data mining program that can analyze the data using ML techniques and can be done by anybody without coding skills.

Overall Assessment of paper

Yes, it is straightforward to read, and this review article is focused on how data is changing with rising technologies and how the complexity of bio medical data is expanding everyday, making it tough to evaluate the data. SIMON (Sequential Iterative Modeling "Over Night"), which analyzes complicated bio medical data, is one such tool designed by the author that can readily evaluate heterogeneous data.

**Research Methodology of paper:**

The author explains how biomedical researchers struggle to evaluate and visualize complex heterogeneous data in this paper's study technique. The author then creates SIMON (Sequential Iterative Modeling "Over Night"), an open-source tool that uses 180+ AI-ML algorithms to readily display and comprehend data as well as analyze very complicated heterogeneous data.

**Future Research**

We need to concentrate on how data is evolving now and how to collect huge and complicated data, which makes data recognition challenging. I also focus on open-source software, such as SIMON, which has 180+ machine learning algorithms for pattern recognition of data and data processing, as well as an open-source software that can be accessed by anyone from anywhere for free and used without any coding knowledge, which is very useful for problem solving.

**New Knowledge Learned:**

# This taught me about how data is evolving nowadays and how to capture vast and complicated data sets that are tough to detect. I also learned about SIMON, an open-source software with 180+ machine learning algorithms for data pattern recognition and processing, as well as an open-source software that can be accessed by anyone from anywhere for free and used without any coding knowledge, which is very useful for problem solving.

# Question to discussed

1. How to design and include the SIMON-required algorithm in line with user demands.

2. How may non-biomedical data be included into the SIMON program?