Cloud-Based Data Preparation for Medical Observational Research

Sathyan, Varun

The future of healthcare lies in endeavors such as precision medicine where treatment options and clinical decisions are tailored to patients based on their exact characteristics, and accurate preventive techniques are implemented. For these advancements to be made possible, clinical data collected from the health records of patients' in healthcare facilities must be provided for analytics, namely observational research, allowing for medical breakthroughs to be made in the treatments of diseases and disorders, widely-held misconceptions to be realized, and important trends to be highlighted. Electronic health records (EHRs) are currently the optimal source of data for these analytics, but, unfortunately, current practices for retrieving relevant EHRs lack automation and efficiency. Clinical data mining is a complex endeavor due to the lack of standardization and structure in the natural text comprising health records. In addition, stringent patient privacy requirements at various organizational levels further complicate procedures for preparing data for observational studies. Based on the characteristics of this issue, with a consideration for the long-term sustainability of the approach, an autonomous, decentralized architecture incorporating a unique application of a semantic similarity algorithm is proposed for automated EHR retrieval, and thus the advancement of personalized medicine.

Awards Won:

Oracle Academy: Award of $5,000 for outstanding project in the systems software category.