



PROJECT TITLE: Data-driven analytics and machine learning for improving healthcare outcome

PROPOSAL NUMBER: Tech3

PI: Lee (GT), Mastrangelo (UW), Tucker (PSU)

RESEARCH THEME: Analytics & Innovative Technologies

BUDGET: \$150,000

MULTI-UNIVERSITY PROJECT: YES

PROJECT YEAR: 1

DESCRIPTION:

Data-driven healthcare has the potential to revolutionize care delivery and trim costs. A major challenge is that providers must sift through and analyze mountains of disparate data to materialize the substantial gain. We continue our healthcare innovation through systems and data analytics. Utilizing EMR and various procedural and personal health data, along with social and behavioral information, we will address all aims - with specific regard to radiologic exam variability. This also had implications in utilizing predictive models to use at the point-of-care when treating infectious disease.

HOW THIS IS DIFFERENT THAN RELATED RESEARCH:

This is the first study where 1) large amounts of patient data are extracted unbiased and globally analyzed, 2) automated encryption of PHI and data integration through terminology mapping is achieved using natural language processing, 3) time series clustering is done with consideration of disease progression despite sparse and missing data, 4) discriminatory factors that inform key decisions are systematically selected using machine learning, 5) individual patient conditions are addressed with the design of personalized evidence-based treatment methods, and 6) this research will be able to be replicated to other cases and sites to improve their process.

EXPERIMENTAL PLAN:

1. Conduct & benchmark literature review
2. Data collection, extraction, and encryption of PHI
3. Data cleaning and integration across types of records and multiple sites
4. Build predictive models using machine learning and derive knowledge for personalized treatment, resource utilization and treatment procedure optimization for best outcome
5. Run pilot projects with optimized procedures and analyze preliminary findings to further refine processes

EXPECTED MILESTONES:

- Objective 1: Use MRI log file data to identify variability and "wasted" time opportunities and to develop predictive models of exam duration, idle time, and repeated scans
- Objective 2: Leverage the size and availability of population health data to model and predict machine utilization efficiency
- Objective 3: Apply machine learning techniques to electronic Health Records containing patient demographics, labs, medications, procedures, and clinical notes to establish health trends and uncover definitive factors that can predict treatment outcome and optimal patient care characteristics
- Objective 4: Design evidence-based expert decision support system to facilitate early diagnosis, optimize and personalize treatment, and ensure safety and reduce errors to provide high quality outcome

BENEFITS TO INDUSTRY:

1) Industry practitioners can make more informed decisions and achieve care that is personalized, timely, evidence-based, and appropriate; 2) Optimize the usage of hospital resources, treatment process, and outcome; 3) Reduce waste, risk, and cost associated with procedures and operations; 4) This research can be replicated to other cases and sites to improve process.

EXPECTED DELIVERABLES:

1) Develop predictive models for KPI's (Exam Duration, Idle Time, Ratio of Repeated Scans), using logfile variables; 2) Define better sequence of scans per exam (exam cards); 3) Design pilot project improving exam cards; 4) Develop models that predict machine utilization efficiency; 5) Design evidence-based expert decision support system and optimize personalized treatment plans.