The Future of Athlete Care – *Diagnostic Support thru Machine Learning*

31st Annual Association of Boxing Commissions Conference, Scottsdale Resort at McCormick Ranch

Association of Boxing Commissions and Combative Sports

Michael Bergeron, Ph.D., FACSM

Senior Vice President

SIVOTEC Analytics

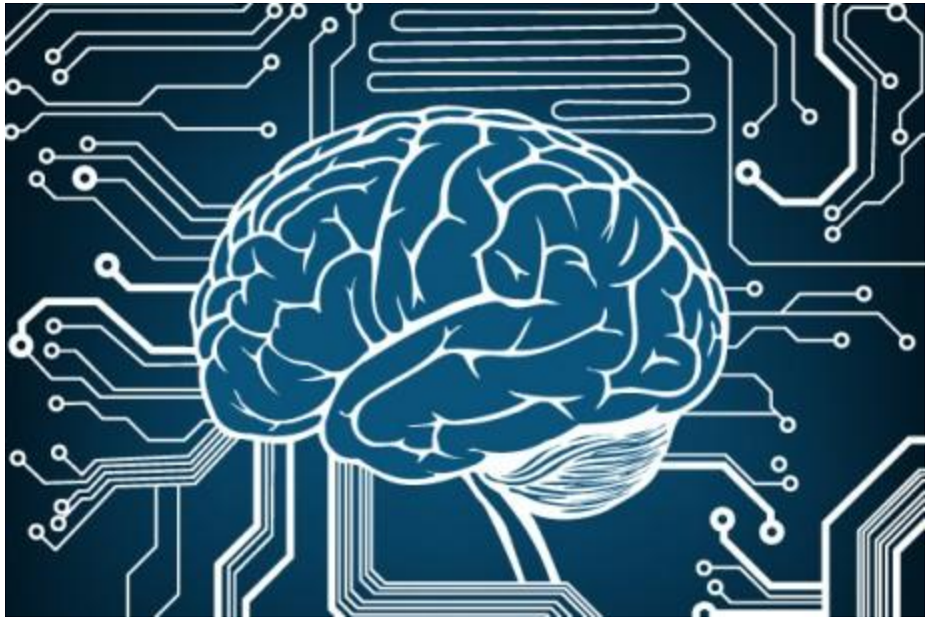
Boca Raton Innovation Campus

Boca Raton, Florida

SIVOTEC ANALYTICS

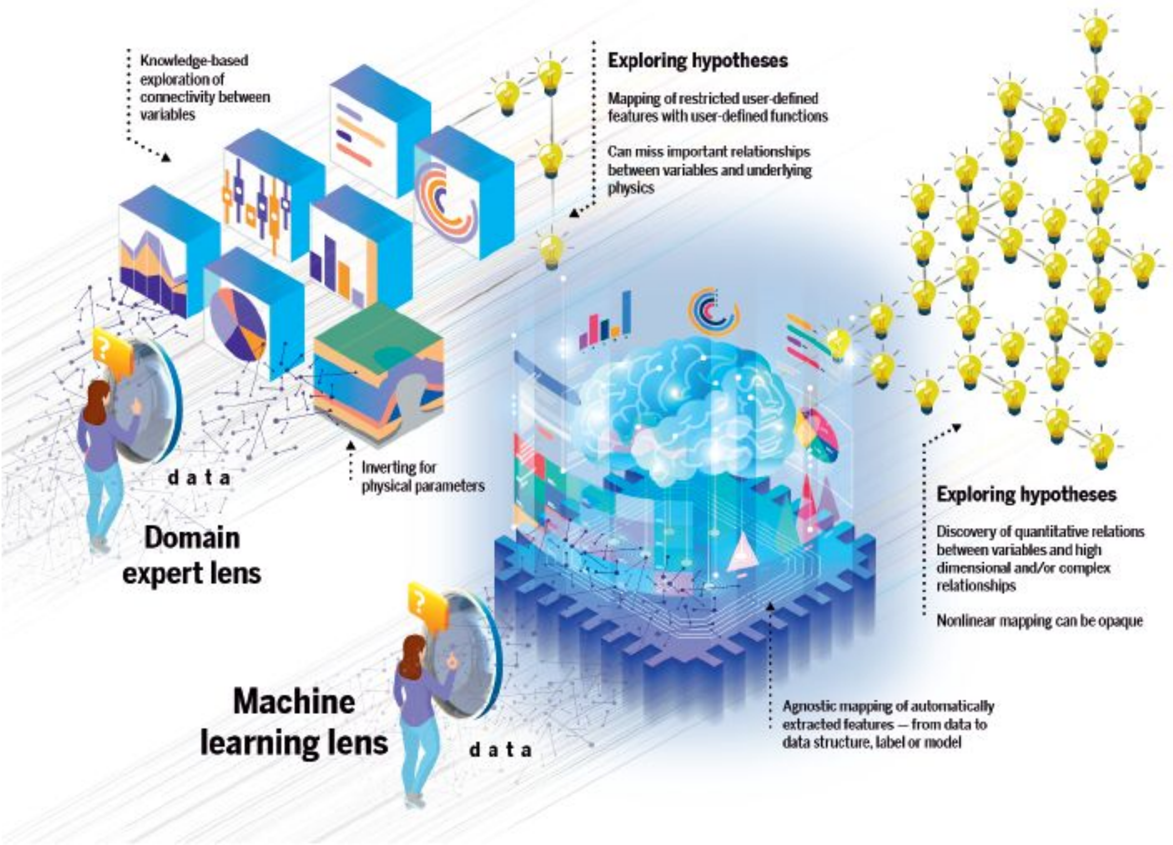
Concussion & Brain Health

- Complex Integrated Systems
  - Anatomy
  - Physiology
  - Functional Neurology
  - Psychology & Behavior
  - Sociology
- Neurocognitive tests
- Balance, vision, biomarkers
- Unstructured – e.g., voice
- Interdependency
- *Synthesize? Interpret? Track?*



Discrete measures comparisons don't tell the whole story

Real-world Complex Systems Biology



- Artificial intelligence, supervised & unsupervised machine learning, super computers
- Machine learning – *best with extensive & high-dimensional data*
- No prior assumptions about distribution or relationships
- Identify notable patterns and high-value features – agnostic (unbiased) mapping
- Build and validate real-world models and decision-support applications
- Incorporate and learn from new data

Questions in *Managing* Brain Health…

Sport Concussion

- Protracted Recovery?
- Conservative Approach?
- Athlete Needs and Accommodations?
- Questions: *Coaches, Parents, Other Stakeholders*
- Expectations?

Degenerative Brain Disease

- Early Detection?
- High-value Indicators?
- Testing
  - Age-specific Norms
  - Sensitivity to Changes?
  - Easy to Use/Engaging?
  - Repeatable and Valid?
- Complement Other Indicators?

Machine learning and predictive modeling – Brain health stratification

Machine Learning in Modeling High School Sport Concussion Symptom Resolve

MICHAEL F. BERGERON[1], SARA LANDSET[2], TODD A. MAUGANS[3], VERNON B. WILLIAMS[4], CHRISTY L. COLLINS[5], ERIN B. WASSERMAN[5], and TAGHI M. KHOSHGOFTAAR[2]

[1]*SIVOTEC Analytics, Boca Raton, FL;* [2]*Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL;* [3]*Division of Neurosurgery, Nemours Children's Hospital, Orlando, FL;* [4]*Kerlan-Jobe Center for Sports Neurology, Los Angeles, CA; and* [5]*Datalys Center for Sports Injury Research and Prevention, Inc., Indianapolis, IN*

ABSTRACT

BERGERON, M. F., S. LANDSET, T. A. MAUGANS, V. B. WILLIAMS, C. L. COLLINS, E. B. WASSERMAN, and T. M. KHOSHGOFTAAR. Machine Learning in Modeling High School Sport Concussion Symptom Resolve. *Med. Sci. Sports Exerc.*, Vol. 51, No. 7, pp. 1362–1371, 2019. **Introduction:** Concussion prevalence in sport is well recognized, so too is the challenge of clinical and return-to-play management for an injury with an inherent indeterminant time course of resolve. A clear, valid insight into the anticipated resolution time could assist in planning treatment intervention. **Purpose:** This study implemented a supervised machine learning–based approach in modeling estimated symptom resolve time in high school athletes who incurred a concussion during sport activity. **Methods:** We examined the efficacy of 10 classification algorithms using machine learning for the prediction of symptom resolution time (within 7, 14, or 28 d), with a data set representing 3 yr of concussions suffered by high school student-athletes in football (most concussion incidents) and other contact sports. **Results:** The most prevalent sport-related concussion reported symptom was headache (94.9%), followed by dizziness (74.3%) and difficulty concentrating (61.1%). For all three category thresholds of predicted symptom resolution time, single-factor ANOVA revealed statistically significant performance differences across the 10 classification models for all learners at a 95% confidence interval ($P = 0.000$). Naïve Bayes and Random Forest with either 100 or 500 trees were the top-performing learners with an area under the receiver operating characteristic curve performance ranging between 0.656 and 0.742 (0.0–1.0 scale). **Conclusions:** Considering the limitations of these data specific to symptom presentation and resolve, supervised machine learning demonstrated efficacy, while warranting further exploration, in developing symptom-based prediction models for practical estimation of sport-related concussion recovery in enhancing clinical decision support. **Key Words:** ADOLESCENT, AUGMENTED INTELLIGENCE, RECOVERY, SPORTS MEDICINE, TRAUMATIC BRAIN INJURY

Concussion prevalence in sport is well recognized, so too is the challenge of clinical and return-to-play management for an injury with an inherent indeterminant time course of resolve (1–7). Although it has been commonly reported that most athletes generally recover from a sport-related concussion (SRC) in 7–10 d postinjury (8), a more prolonged recovery course is increasingly recognized and evident for many of those affected (4,9,10). Notably, the most recent Berlin consensus statement on concussion in sport notes that the term "persistent symptoms" after SRC should be used when symptoms linger beyond the normal clinical recovery time frames (i.e., >10–14 d in adults and >4 wk in children) (4). However, regardless of the criteria for prolonged recovery, clear and valid insight into the anticipated resolution time based on reported symptoms and other available relevant information could measurably assist in planning an individualized stratified care approach to medically managing SRC in young athletes.

Certain SRC symptoms are more characteristically prevalent at initial presentation (e.g., headache, fatigue, and dizziness), whereas others (e.g., sleep disturbance, frustration, and forgetfulness) typically develop subsequently as delayed or unclear timing of onset and/or reporting, and each has varying durations through the course of concussion recovery (9,10). Several notable factors have been reported to increase the risk for prolonged SRC symptoms in youth, including a history of recent or multiple concussions (1,11), although these findings are inconsistent (10,12). Young age has also been recognized as a modulating factor specific to reported symptoms and respective severity and duration, although the effect may be only small (13). Some evidence indicates that girls report more symptoms and are purportedly more susceptible to a protracted recovery course than boys (14). However, other findings do not support this perspective (10).

Previous approaches have examined selected symptoms in attempt to clarify an anticipated SRC protracted recovery.

Address for correspondence: Michael F. Bergeron, Ph.D., F.A.C.S.M., SIVOTEC Analytics, Boca Raton Innovation Campus, 4800 T-Rex Avenue, Suite 315 Boca Raton, FL 33431; E-mail: mbergeron@sivotecanalytics.com.
Submitted for publication July 2018.
Accepted for publication January 2019.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/19/5107-1362/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000001903

Journal of Alzheimer's Disease xx (20xx) x–xx
DOI 10.3233/JAD-190165
IOS Press

Episodic-Memory Performance in Machine Learning Modeling for Predicting Cognitive Health Status Classification

Michael F. Bergeron[a,*], Sara Landset[b], Franck Tarpin-Bernard[c], Curtis B. Ashford[d], Taghi M. Khoshgoftaar[b] and J. Wesson Ashford[e]
[a]*SIVOTEC Analytics, Boca Raton, FL, USA*
[b]*Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA*
[c]*HAPPYneuron, S.A.S., Lyon, France*
[d]*MemTrax, LLC., Redwood City, CA, USA*
[e]*War-Related Illness and Injury Study Center, VA Palo Alto Health Care System and Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Palo Alto, CA, USA*

Handling Associate Editor: David Loewenstein

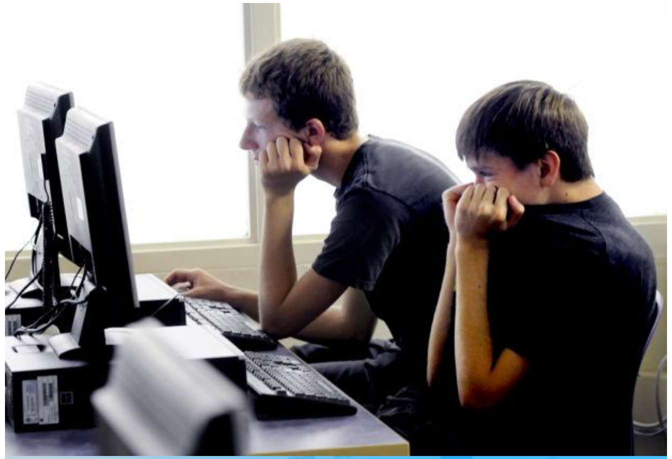
Accepted 1 May 2019

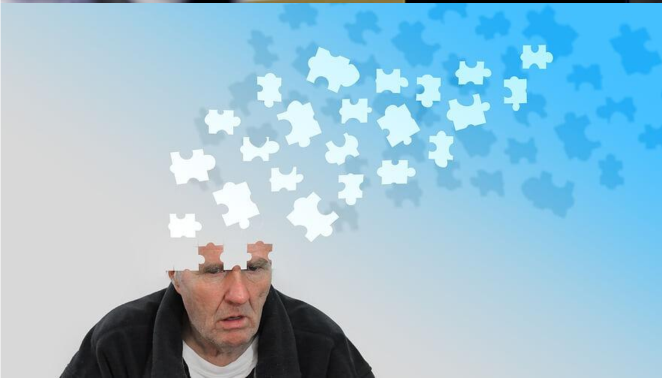
Abstract.
Background: Memory dysfunction is characteristic of aging and often attributed to Alzheimer's disease (AD). An easily administered tool for preliminary assessment of memory function and early AD detection would be integral in improving patient management.
Objective: Our primary aim was to utilize machine learning in determining initial viable models to serve as complementary instruments in demonstrating efficacy of the MemTrax online Continuous Recognition Tasks (M-CRT) test for episodic-memory screening and assessing cognitive impairment.
Methods: We used an existing dataset subset ($n = 18{,}395$) of demographic information, general health screening questions (addressing memory, sleep quality, medications, and medical conditions affecting thinking), and test results from a convenience sample of adults who took the M-CRT test. M-CRT performance and participant features were used as independent attributes: true positive/negative, percent responses/correct, response time, age, sex, and recent alcohol consumption. For predictive modeling, we used demographic information and test scores to predict binary classification of the health-related questions (yes/no) and general health status (healthy/unhealthy), based on the screening questions.
Results: ANOVA revealed significant differences among HealthQScore groups for response time true positive ($p = 0.000$) and true positive ($p = 0.020$), but none for true negative ($p = 0.0551$). Both %responses and %correct had significant differences ($p = 0.026$ and $p = 0.037$, respectively). Logistic regression was generally the top-performing learner with moderately robust prediction performance (AUC) for HealthQScore (0.648–0.680) and selected general health questions (0.713–0.769).
Conclusion: Our novel application of supervised machine learning and predictive modeling helps to demonstrate and validate cross-sectional utility of MemTrax in assessing early-stage cognitive impairment and general screening for AD.

Keywords: Aging, Alzheimer's disease, dementia, mass screening

*Correspondence to: Michael F. Bergeron, PhD, FACSM, SIVOTEC Analytics, Boca Raton Innovation Campus, 4800 T-Rex Avenue, Suite 315, Boca Raton, FL 33431, USA. E-mail: mbergeron@sivotecanalytics.com.

ISSN 1387-2877/19/$35.00





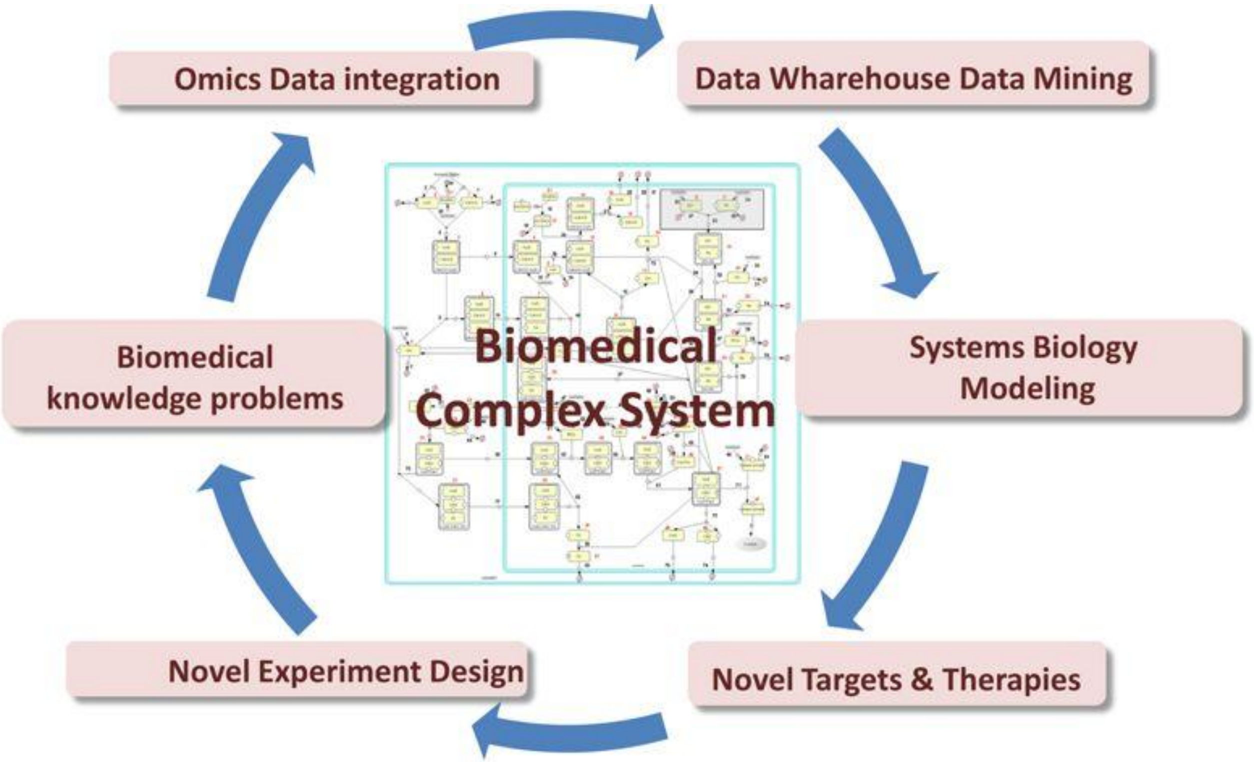
Technology Transfer – *Practical Clinical Application*

- Technology transfer usable tools
  - Aim – *Effective prediction models*
  - Dimensionality (spatial data map) reduction/simplification
  - Feature ranking/engineering
  - Top-performing classifiers
  - Embedded in the application
  - Utility and new information (new cases) improves the models – inductive *learning*

Benefits of Informed Stratified Classification

- Augment/prioritize *stratified patient care* vs. step-wise after “failure to respond”
- *Individualized* care – manage expectations
- Coaches, family, et al. – *planning accommodations*
- Mitigate returning too soon
- Research benefits – subject selection

Tomorrow’s Models



The Future? Already Here!

- 21st Century Technology
  - Complementary Instrument
  - Diagnostic/Decision <u>Support</u>
  - Anticipate/Stratify/Prioritize Patient Care
  - Reveals Individual Solutions
  - Trends Alerts in Real-Time
  - Makes You Better!
  - *Helps the Athletes…*