BCIL
Biomedical & Clinical Informatics Lab

# Application of a Tensor-Based Classification Method with Electrocardiogram Data

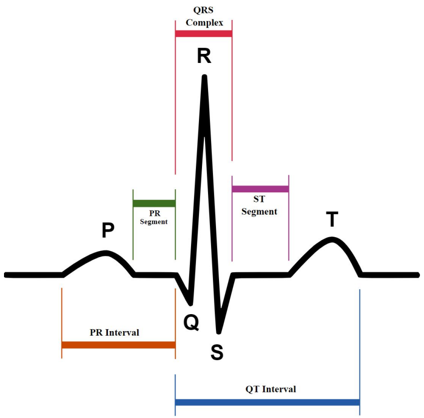
Olivia Alge, Cristian Minoccheri, Jonathan Gryak, Harm Derksen, J. Scott VanEpps, Kayvan Najarian

BCIL
Biomedical & Clinical Informatics Lab

# Introduction

# Electrocardiogram (ECG)

- Abbreviated as “ECG” or “EKG”
- Records electrical activity of the heart using electrodes
  - Generally, 10 electrodes are placed on the body to generate 12 leads




Original image created by Agateller (Anthony Atkielski), converted to svg by atom. (https://commons.wikimedia.org/wiki/File:SinusRhythmLabels.svg), “SinusRhythmLabels“, background color modified

BCIL
Biomedical & Clinical Informatics Lab

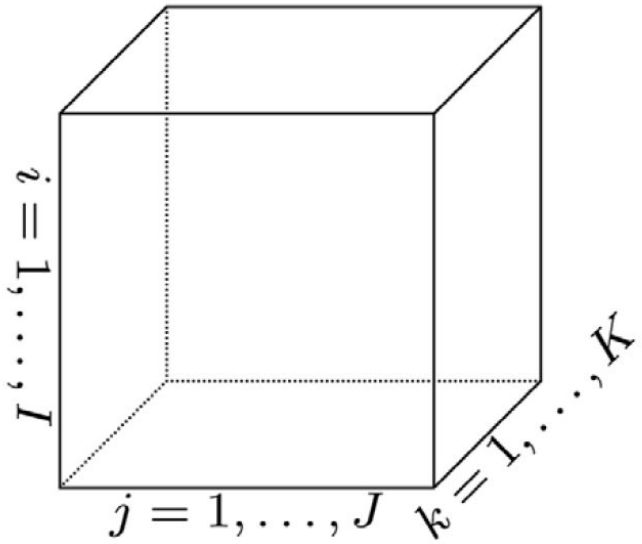
# Sepsis

- Infection-induced syndrome with life-threatening organ dysfunction [5]
  - Increase in Sequential Organ Failure Assessment (SOFA) score of 2 or more points
    - Criteria based on statuses of:
      - Respiration
      - Coagulation
      - Liver
      - Cardiovascular system
      - Renal function
      - Central nervous system
    - qSOFA for bedside care
- Septic Shock
  - Subset of sepsis with profound circular, cellular, and metabolic abnormalities and higher risk of mortality
    - Despite adequate fluid resuscitation, vasopressors needed to maintain Mean Arterial Pressure >= 65mm Hg, and serum lacatate level > 2 mmol/L [5]

# Sepsis

- Risk stratification for sepsis-induced organ dysfunction is difficult because of its heterogeneity.
- Information to compute SOFA may not be immediately available for patients that require care [5].
- Continuous ECG recording is often available for patients in intensive care
  - May be useful to identify those at risk of decompensating to septic shock
  - Structuring multi-lead ECG data as a higher-order tensor can:
    - Preserve the data's structure
    - Prevent loss of information that can occur in traditional machine learning

# Tensors




A third-order tensor: $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$

Image adapted from [3]

Biomedical & Clinical Informatics Lab

# Tensor Analysis

- Why use tensors?
  - Traditional learning methods (RF, SVM, LDA) flatten information into vectors
    - Loss of information, structure
    - Difficulty seeing interactions among features with respect to time
- Therefore, represent heterogeneous data in sepsis dataset as a tensor
  - Prevent loss of information
  - Preserve structure
  - View changes and interactions among features over time

# Goals

- The project discussed here is a first step in this research:
  - Development of a tensorial data classification method
  - Testing the method using publicly available ECG data

Biomedical & Clinical Informatics Lab

# Generating Tensors

- Use publicly available Physionet PTB database [1,4]
  - 12-lead ECG recordings from 216 unhealthy subjects, 52 healthy controls
    - Unhealthy cases include myocardial infarction, cardiomyopathy, dysrhythmia, etc.
- Process each 12-lead sample
  - Select signals >= 90 seconds
  - Crop ECG signals to 90 seconds in length
    - 129 unhealthy, 24 healthy
  - Remove noise and baseline wander with Butterworth filter
  - Add specified level of Gaussian noise
  - Apply Taut String [2] with 5 specified parameter values
  - Extract 6 features from taut string approximation
  - Yield 12 x 5 x 6 tensor

# Tensor Classification:
# Kempf-Ness Multilinear Discriminant Analysis (KNMDA)

- For each class, solve the minimization problem:

$$\underset{(A,B,C)}{\operatorname{argmin}}\left(\sum_{i=1}^{N_1} \|(A, B, C) \cdot \mathcal{X}_i\|^2\right)$$

- We can interpret group action via (A, B, C) as a change of coordinates for centered tensors in a given class
  - Iterate solving for A, B, C until the norm is reduced beyond a specified tolerance level, or a maximum number of iterations is reached
- To perform classification on a new tensor Z of unknown class:
  - Compute $d = \left\|(A, B, C) \cdot (\mathcal{Z} - \bar{\mathcal{X}})\right\|$ for each class
  - Assign Z to the class that generates the smallest distance

# Classification of Tensors Extracted from ECG

| Method | Clean | SNR = -10 | SNR = -30 |
|---|---|---|---|
| **KNMDA** | 0.94 (0.03) | 0.92 (0.03) | 0.88 (0.04) |
| **DATEReig** | 0.90 (0.04) | 0.81 (0.06) | 0.80 (0.06) |
| **DATER** | 0.90 (0.04) | 0.83 (0.05) | 0.82 (0.04) |
| **CMDA** | 0.91 (0.03) | 0.80 (0.08) | 0.79 (0.07) |
| **ManTDA** | 0.87 (0.05) | 0.72 (0.10) | 0.70 (0.12) |
| **SVM** | 0.91 (0.03) | 0.85 (0.06) | 0.85 (0.05) |
| **LDA** | 0.68 (0.07) | 0.59 (0.07) | 0.56 (0.07) |

# Results

Over thirty iterations, our KNMDA method obtained average area under the receiver operating characteristic curve (AUC) of 0.94 (SD 0.03), an improvement compared to linear SVM with 0.91 (0.03) and LDA with 0.68 (0.07) with no additional noise. When signal-to-noise ratio of the ECG signal was changed to -10, KNMDA achieved average AUC of 0.92 (0.03), compared to SVM with 0.85 (0.06) and LDA with 0.59 (0.07).

# Impact

The results show that our proposed method can outperform existing ones when using real ECG data. We can therefore apply this method to our larger goal of improving risk-stratification for septic patients, to augment physician decision-making and improve patient care.

# References

[1] Bousseljot R, Kreiseler D, Schnabel, A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomedizinische Technik, Band 40, Ergänzungsband 1 (1995) S 317

[2] P. L. Davies and A. Kovac, "Local extremes, runs, strings and multiresolution," *Annals of Statistics*, pp. 1–48, 2001.

[3] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, Aug. 2009, doi: 10.1137/07070111X.

[4] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.

[5] M. Singer *et al.*, "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)," *JAMA*, vol. 315, no. 8, pp. 801–810, Feb. 2016, doi: 10.1001/jama.2016.0287.

# Funding


- Supported by the University of Michigan NIH NIGMS Bioinformatics Training Grant of the National Institutes of Health under award number T32GM070449
- The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

# Acknowledgments

- Biomedical and Clinical Informatics Lab (BCIL)
    - Maryam Bagherian
    - Harm Derksen
    - Jonathan Gryak
    - Gang Liu
    - Cristian Minoccheri
    - Kayvan Najarian
    - S.M.Reza Soroushmehr
- VanEpps Lab
    - J. Scott VanEpps

BCIL
Biomedical & Clinical Informatics Lab

# Thank you