COPENHAGEN

47TH ANNUAL

International Conference of the

IEEE Engineering in Medicine and Biology Society

AI-Based Performance Analysis for Track and Field Athletes

Mostafa Habibi, Mehrdad Nourani

Department of Electrical & Computer Engineering

The University of Texas at Dallas

Richardson, Texas 75080, USA

Mohammad Mehdi Nourani

Athletic Department

California Polytechnic State University

San Luis Obispo, CA 93407, USA

Outline

- Background & Motivation
  - Data-Driven Analytical Need & Solution
- Prior Works
- Proposed Methodology
  - Data Collection & Curation
  - Feature Extraction & Transformation
  - Online Clustering
- Experimental Results
  - Cluster separation
  - Internal & External validation
  - Performance ranking & prediction
- Conclusion and Future Work

Motivation

- Accurate evaluation of athletes' performances presents a significant challenge for coaches

- Traditional methods rely on personal best times (PBs), which may not reflect true athletic ability
  - A runner with one exceptional time but inconsistent results may be overrated.
  - Consistently solid racers, even without records, often show higher long-term potential

- The versatility is often overlooked.
  - Coaches often overlook how well runners adapt to different events or conditions

Data-Driven Analytical Need & Solution

- Currently, no standard data-driven method exists to compare athletes or evaluate athletes' performance trends over time
- Platforms like TFRRS provide raw competition data but lack advanced analytics
- Coaches struggle to identify trends, predict outcomes, and rank athletes by consistency, versatility, and improvement.
- Proposed solution
  - Approach integrating feature extraction, clustering, and trend analysis
  - Helping coaches give better feedback and more personalized training plans

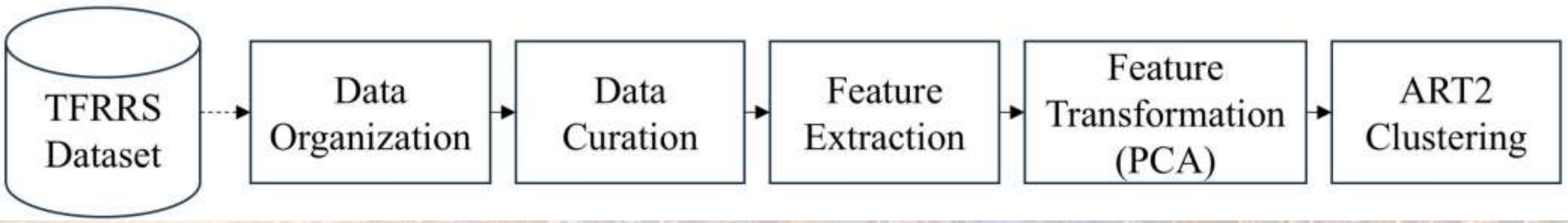
Prior Works

- Evaluation and prediction of athletic performance rely on data from four key domains:
  - **Anthropometric**: Body mass and height (impact speed, power, endurance) [J. Gómez-Molina et al., Journal of Sports Science & Medicine, 2017]
  - **Training**: Weekly running volume, intensity, frequency, and years of experience (show preparation and correlation to performance) [P. Chatzakis et al., Journal of Physical Education, 2019]
  - **Physiological**: VO_2 max, maximum velocity, anaerobic threshold (reflect aerobic/anaerobic capacities) [D. Blythe & F. Király, PLOS ONE, 2016]
  - **Biomechanical**: Step length, step rate, ground contact time (evaluate running mechanics) [M. Habibi et al., IEEE SMARTCOMP, 2024]

- Limitations of Existing Studies
  - Most studies ignore historical performance trends and consistency.
  - No known research has correlated historical data to cluster/rank collegiate or new athletes.
  - Addressing this can improve training strategies and competitive outcomes.

Proposed Methodology

- Our approach uses advanced machine learning, statistical analysis to analyze runners' historical performance data from TFRRS.



- It provides coaches with explainable insights into athletes' consistency, improvement trends, predicted rankings, and personalized training decisions.

Step 1: Data Collection

- We collected data from TFRRS, a comprehensive public domain platform to report collegiate Track and Field performances.
  - Established in 2010 as a mandatory system for NCAA member institutions
  - It provides detailed records of competitions, athletes, dates and environmental factors
    - Example of records:

Athlete ID	Year	Team	Field	Types	Score	Competition venue	Date	Wind
1MT	FR-1	Tennessee	200 Meters	Outdoor	20.56	NCAA East First Round	05/22/2024	-0.5
2AS	SO-1	South Florida	200 Meters	Indoor	21.24	American Indoor Track & Field Championship	02/23/2024	-
3OG	JR-3	LSU	200 Meters	Outdoor	20.74	LSU Battle on the Bayou	08/29/2024	2.1
4MK	SR-4	Alabama	400 Meters	Outdoor	45.76	NCAA I Outdoor Track & Field Championships	06/05/2024	0

Data Collection → Data Curation → Feature Extraction → Feature Transformation → ART2 Clustering

Step 1: Data Collection – Cont.

- We structured TFRRS data into tables, extracting athlete info, performance, venue, and wind speed for analysis.
- Eight datasets of top 500 NCAA Division I men and women in 200m and 400m, segmented by gender, field, and type, were created for detailed trend analysis.

Field	Gender	Type	# of Scores
200 Meters	Women	Indoor	4137
		Outdoor	5075
	Men	Indoor	4821
		Outdoor	5461
400 Meters	Women	Indoor	3107
		Outdoor	3833
	Men	Indoor	4211
		Outdoor	4541

Data Collection → Data Curation → Feature Extraction → Feature Transformation → ART2 Clustering

Step 2: Data Curation

- Three-step cleaning process
  1. **Removing Outlier Scores**
     - Removed scores 10% above or below a runner's average, often due to injury or same-day errors.
  2. **Filtering Runners with Insufficient Data**
     - Excluded runners with under 4 scores due to insufficient data for trend analysis.
  3. **Addressing Environmental Factors**
     - Removed outdoor scores with wind beyond ±2.0 m/s (IAAF rules) to ensure fair, valid results.

- This process removes inaccuracies, inconsistencies, and biased scores to have a reliable dataset

Data Collection → Data Curation → Feature Extraction → Feature Transformation → ART2 Clustering

Step 3: Feature Extraction

- Extracting features to derive meaningful, impactful metrics for each runner

1. **Best Score:** Minimum time recorded in a specific field
2. **Average Score:** Mean of all recorded scores in a specific field
3. **Variance Score:** Provide insights into consistency
   - High variance → fluctuating performance (inconsistency)
   - Low variance → steady performance (reliability)
4. **Consistency (CV):** The coefficient of variation provides a normalized measure of variability
   - Lower CV implies greater consistency across athletes or events.
5. **Trend Analysis Using Fast Fourier Transform (FFT)**
   - Reveals hidden patterns & periodic trends
   - Helps understand long-term performance evolution

Data Collection → Data Curation → **Feature Extraction** → Feature Transformation → ART2 Clustering

Step 4: Feature Transformation (Dimension Reduction)

- Principal Component Analysis (PCA) reduces dimensionality by transforming correlated features into uncorrelated principal components.

- This transformation ensures that the new features retain most of the variability in the original dataset while reducing its dimensionality.

- This process involved:
  - Constructing the covariance matrix to capture feature relationships.
  - Using eigenvectors (directions) and eigenvalues (variance) to find the main components.

Data Collection → Data Curation → Feature Extraction → **Feature Transformation** → ART2 Clustering

Step 4: Feature Transformation (Dimension Reduction) – Cont.

- Using PCA

1. The covariance matrix capture how features relate to each other.
   - A new mathematical feature (MF) was defined:

$$MF = w_{11}x_1 + w_{12}x_2 + \cdots + w_{1n}x_n$$

where w_{ij} are coefficients from the covariance matrix.

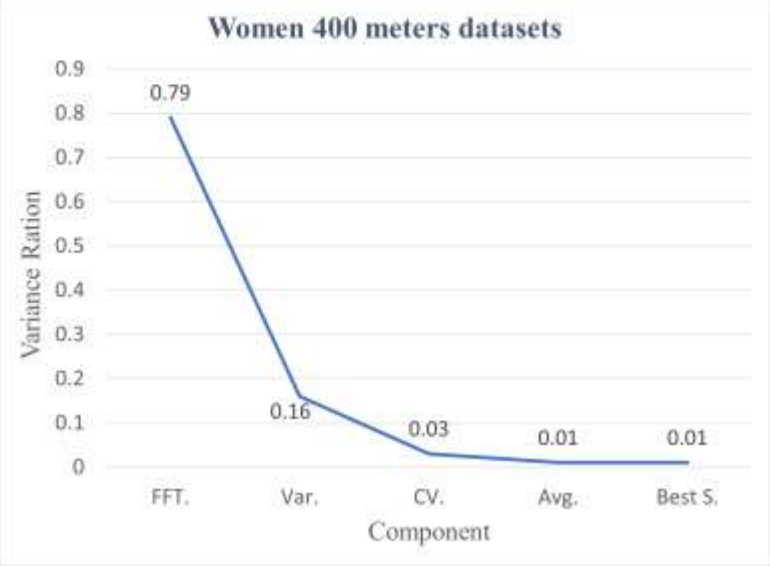
   - This approach leverages the contribution of individual components to represent the dataset.

Data Collection → Data Curation → Feature Extraction → **Feature Transformation** → ART2 Clustering

Step 4: Feature Transformation (Dimension Reduction) – Cont.

- Using PCA

2. The variance ratio tells us which components explain most of the data's variability.
   - The figure shows the variance ratio from women's 400-meter indoor data:
     - FFT contributes 79% of variance
     - Variance contributes 16%.
     - CV, Avg, Best Score combined contribute <5%.



Data Collection
Data Curation
Feature Extraction
Feature Transformation
ART2 Clustering

Step 5: ART2 Clustering

- Adaptive Resonance Theory 2 (ART2) is a clustering algorithm designed for online and adaptive learning
  - ART2 does not require a predefined number of clusters
  - It dynamically adjusts clusters as new data points are added, making it ideal for real-time analysis
- ART2's vigilance parameter (ρ) sets how closely a new data point must match existing clusters
  - If similarity to an existing cluster is above ρ, the point is assigned to that cluster
  - If below ρ, a new cluster is created for it
  - Similarity is measured by the normalized dot product (cosine similarity) between the new data point and the cluster center
- This process is ideal for our application, where new results are added weekly or monthly during the season.

Data Collection → Data Curation → Feature Extraction → Feature Transformation → ART2 Clustering

Experimental Results

- We utilize ART2 algorithm for clustering to
  - Examine runners' performances across 8 datasets
  - Cluster our dataset into 5 clusters (A to E: the best to the lowest-performing runners)

- Analysis Areas:
  - Cluster separation (density distributions)
  - Internal evaluation (Silhouette scores)
  - External validation (box plots)
  - Performance ranking & prediction

Field	Gender	Type	# of Scores	# of Clusters
200 Meters	Women	Indoor	4137	5
		Outdoor	5075	5
	Men	Indoor	4821	5
		Outdoor	5461	5
400 Meters	Women	Indoor	3107	5
		Outdoor	3833	5
	Men	Indoor	4211	5
		Outdoor	4541	5

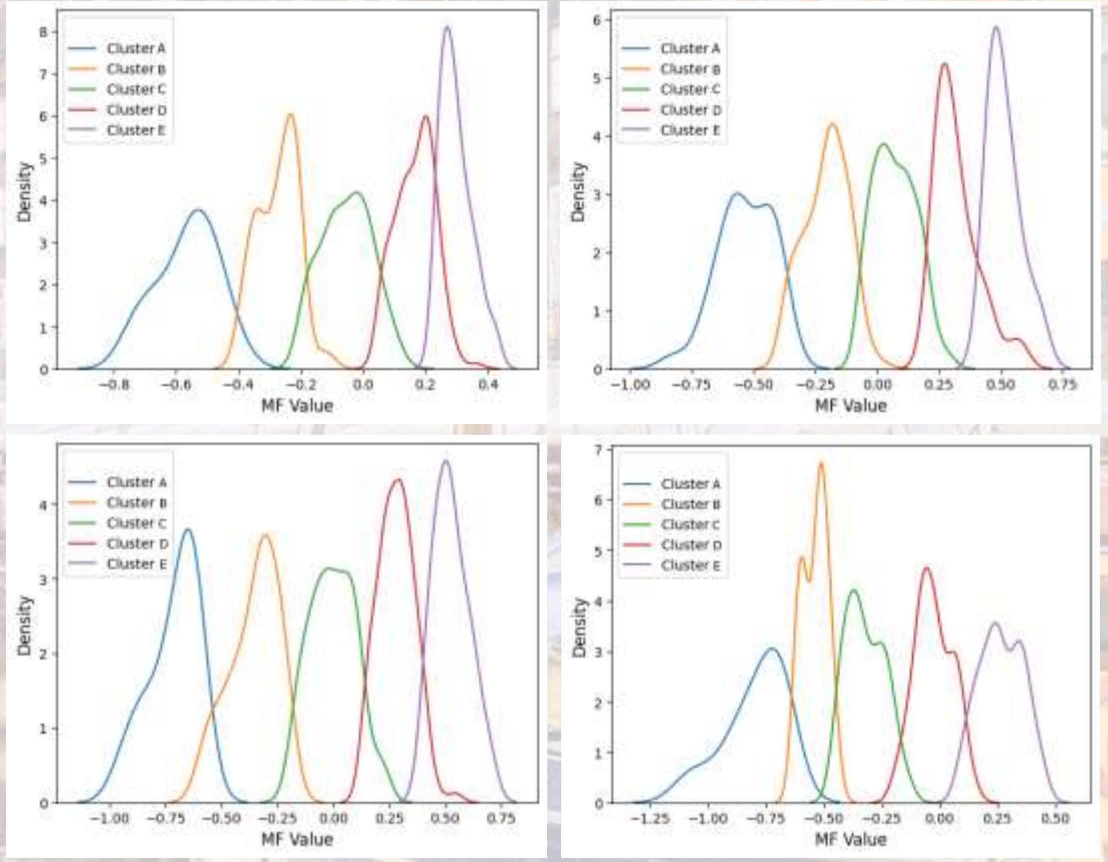
Experimental Results – Cluster Separation

- Probability density distributions of mathematical feature (MF) across the different clusters (A to E) for women datasets.
  - Each curve is the probability density for a cluster; peaks highlight typical MF values.
  - The separation of curves (less overlap) means clusters are well-distinguished by MF.

Indoor Outdoor

200 m

400 m



COPENHAGEN
47TH ANNUAL
International Conference of the
IEEE Engineering in Medicine
and Biology Society
IEEE EMBS

Experimental Results – Internal Evaluation

- We used Silhouette Score for internal evaluation due to a lack of ground truth
  - Measures how well points fit in their clusters vs. others:

$$Silhouette\ score = \frac{b-a}{max\ (a,b)}$$

- Scores close to 1 indicate good separation
  - High scores (~0.89–0.92), confirming effective clustering

Field	Gender	Type	# of Clusters	Silhouette
200 Meters	Women	Indoor	5	0.9
		Outdoor	5	0.92
	Men	Indoor	5	0.91
		Outdoor	5	0.89
400 Meters	Women	Indoor	5	0.89
		Outdoor	5	0.9
	Men	Indoor	5	0.9
		Outdoor	5	0.89

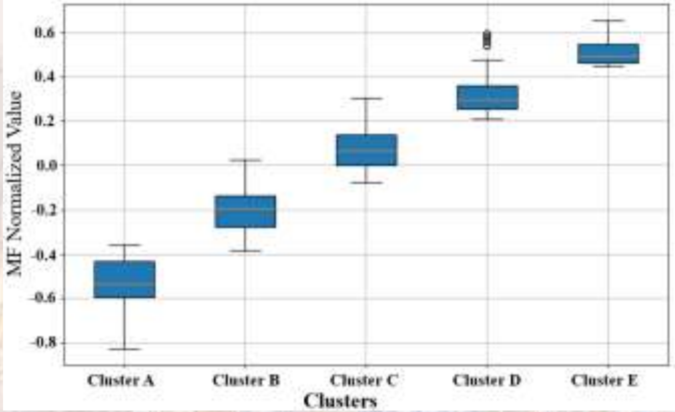
Experimental Results – External Validation

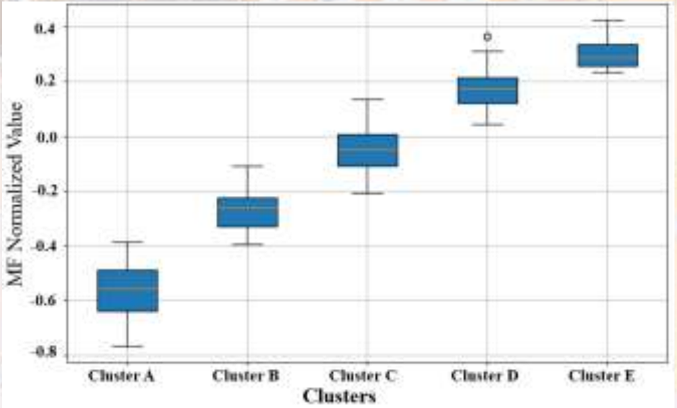
- Figures show clustering results for women's datasets using box plots
  - Shown using box plots to display inter-quartile range (IQR), median, and outliers.
  - Makes it easy to compare cluster ranges and spot variability across groups.
  - Helps quickly assess performance levels and consistency within each cluster.

Indoor

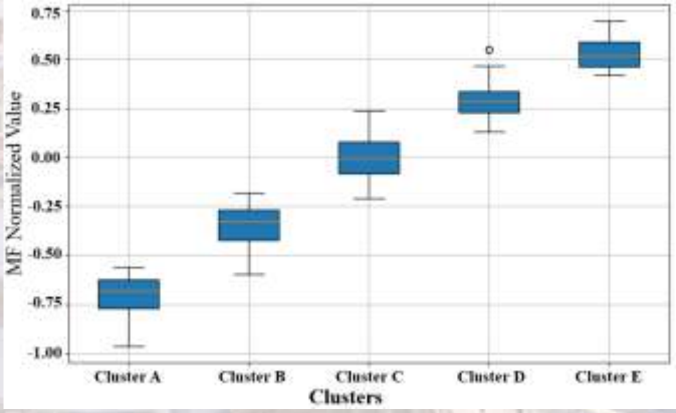
Outdoor

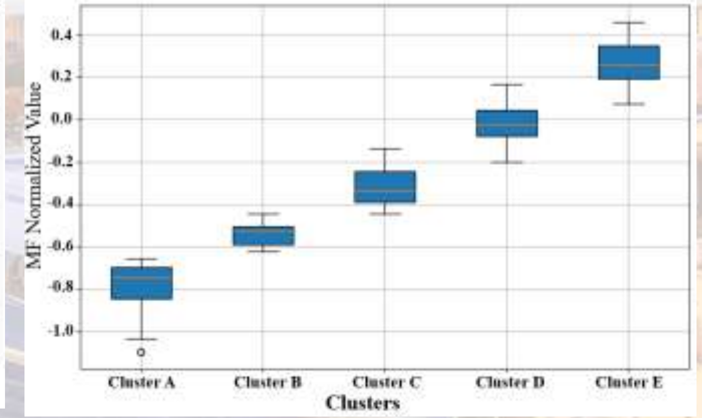
200 m





400 m





Experimental Results – External Validation – Cont.

- The table presents cluster metrics for women's outdoor 200m and 400m races.
- Statistical analysis (based on indoor 200m) shows:
  - Relying only on the best score is inadequate.
  - Using mean, variance, CV, and FFT allows clustering aligned with performance and consistency.

Field	Cluster	Mean Best	Mean Avg	Mean Var	Mean CV	Mean FFT
Outdoor 200 meters	A	22.12	22.86	2.71	1.81	22.6
	B	22.77	23.5	2.52	1.93	23.32
	C	23.27	23.94	3.07	1.86	23.66
	D	23.79	24.39	2.94	1.77	24.16
	E	24.08	24.71	3.75	1.82	24.38
Outdoor 400 meters	A	50.37	51.94	1.32	1.87	51.15
	B	51.69	53.01	1.47	1.59	52.43
	C	52.72	54.12	1.53	1.65	53.48
	D	54.19	55.55	1.35	1.67	54.87
	E	55.66	56.86	1.78	1.49	56.28

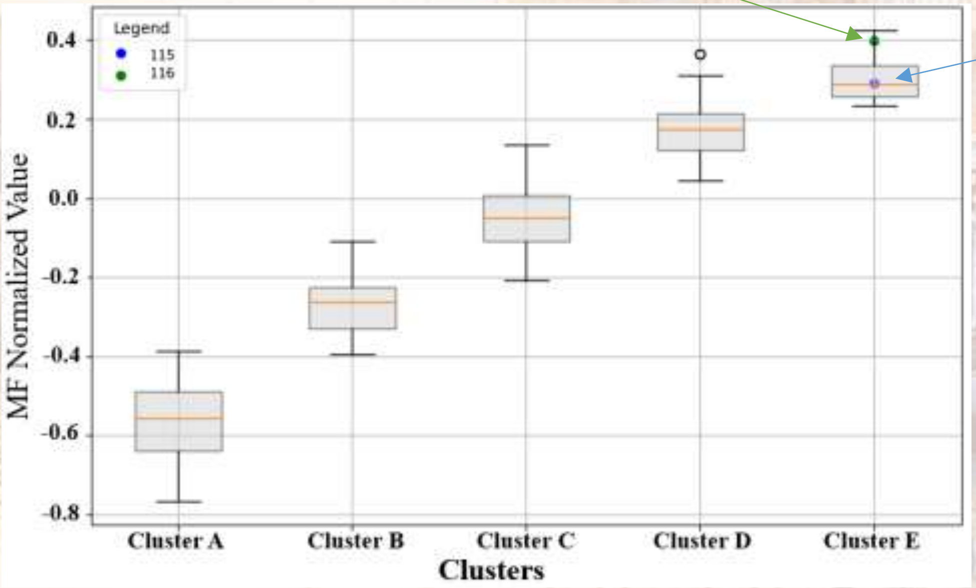
Experimental Results – Performance Ranking & Prediction

- Using historical data from the TFRRS database and clusters of competitive athletes' features, a coach can evaluate and rank a new athlete based on limited recorded performances
- Consider four runners (IDs 113-116) whose historical data is in table
  - Our ranking methodology assessed their performances against Division I race results

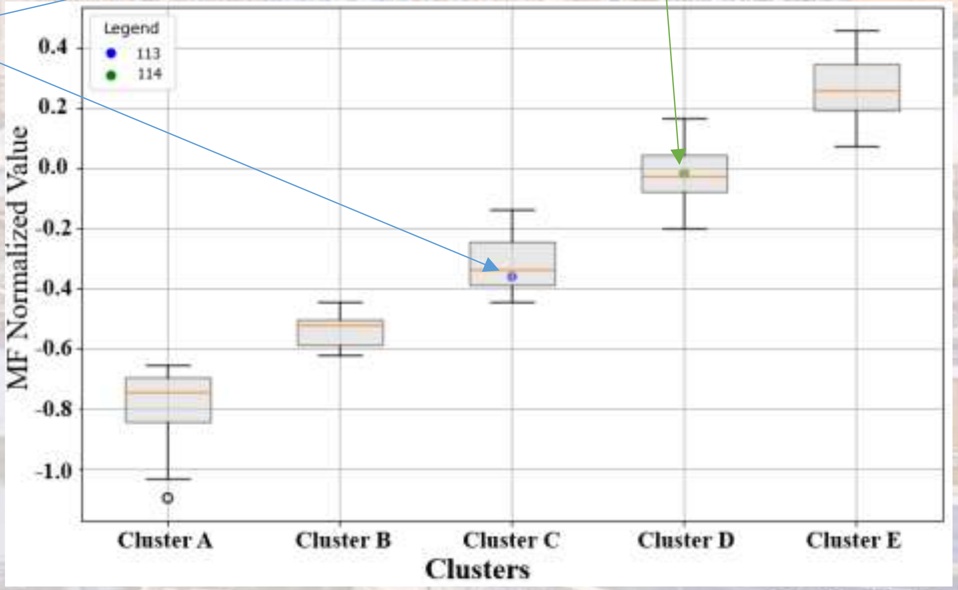
Gender	Field	Type	ID	Training and/or prior	Rank cluster
Women	400m	Outdoor	113	[56.08, 56.61, 51.36, 50.52, 53.81]	C
Women	400m	Outdoor	114	[56.99, 58.59, 53.15, 52.86, 54.01]	D
Women	200m	Outdoor	115	[24.71, 24.85, 25.1, 24.06, 24.8, 24.64, 24.8]	E
Women	200m	Outdoor	116	[25.01, 24.67, 24.8, 24.4, 24.7, 25.1]	E

Experimental Results – Performance Ranking & Prediction – Cont.

- Figures rank new runners against records of past races in Division I
  - In 400 meters, runner 113 (cluster C) is likely to perform better than runner 114 (cluster D).
  - In 200 meters, runner 116 may perform slightly better than runner 115; however, both are in cluster E



Women's outdoor 200m



Women's outdoor 400m

Conclusions

- We proposed a data-driven method in which coaches can assess the performance of their track and field athletes

- Our methodology is very flexible and can be extended to provide a comparative assessment based on training history, competition division, and event type

- This approach enables track & field coaches to
  - Performance Benchmarking
  - Gain insights on consistency and trends
  - Personalized Training
  - Talent Identification
  - Competition Strategy

Thanks for Your Attention!

Mostafa Habibi, Mehrdad Nourani
Department of ECE
The University of Texas at Dallas
{mostafa.habibidehsheikhi, nourani}@utdallas.edu

Mohammad Mehdi Nourani
Athletic Department
California Polytechnic State University
mnourani@calpoly.edu