University of Nebraska at Omaha

UNO BIOINFORMATICS

Mobility and Health: can we use mobility data to accurately monitor health levels and predict potential Monitoring?

IARIA SENSORCOMM 2021

Hesham H. Ali – hali@unomaha.edu
UNO Bioinformatics Core Facility
College of Information Science and Technology

Hesham H. Ali, PhD

Background and Professional Experience:

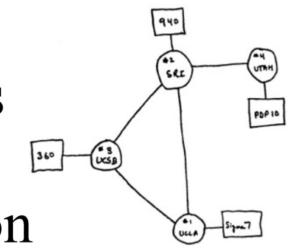
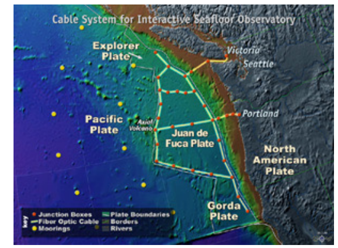
- Professor of Computer Science, College of Information Science and Technology, University of Nebraska Omaha (UNO)
- Director of UNO Bioinformatics Core Facility
- Served as the Lee and Wilma Seemann Distinguished Dean of UNO College of Information Science and Technology between 2006 and 2021
- Published over 200 articles, books and book chapters in various IT areas including scheduling, distributed systems, data analytics, wireless networks, and Bioinformatics
- Has been leading a Research Group that focuses on developing innovative computational tools to analyze all health-related data with the goal of advancing next generation of biomedical research and contributing to the preventive and personalized healthcare initiatives

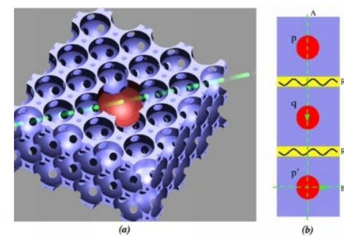
Current Revolution in Scientific Research

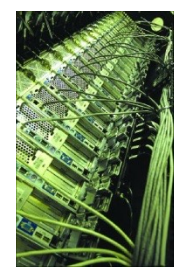
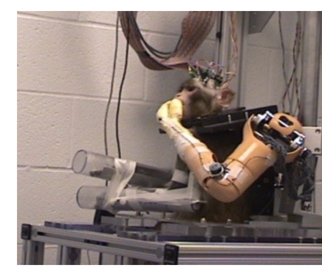
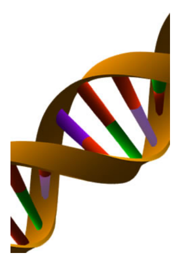
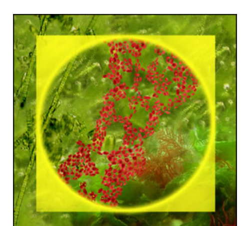
- IT is changing many scientific disciplines
- So much relevant data is currently available
- The availability of data shifted many branches in sciences from pure experimental disciplines to knowledge-based disciplines
- Incorporating Computational Sciences and other branches of sciences is not easy
- Interdisciplinary Research? Translational Research? Big Data Analytics?

The Future is Full of Opportunity

- Creating the future of networking
- Driving advances in all fields of science and engineering
- Revolutionizing transportation
- Personalized education
- The smart grid
- Predictive, preventive, personalized medicine
- Quantum computing
- Empowerment for the developing world
- Personalized health monitoring => quality of life
- Harnessing parallelism
- Synthetic biology

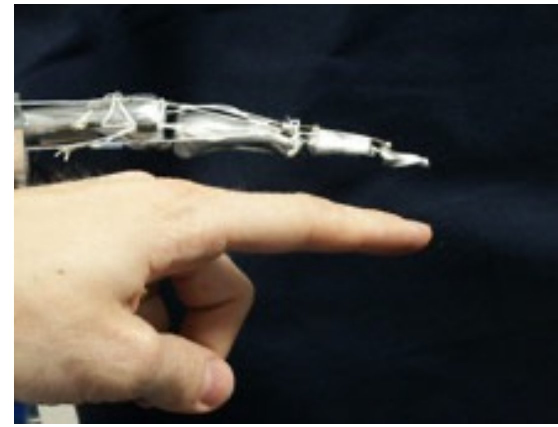







Revolutionizing Health

UNO BIOINFORMATICS

Personalized health monitoring

Evidence-based medicine



Neurobotics

P4 Medicine

INSTITUTE FOR Systems Biology

It's all about the Data!

- How it all began:
  - Advances in instruments and computational technologies led to new new research directions
  - Massive accumulation of data led to investigating new potential discoveries
  - The availability of enormous various types of public/private data sources
  - How to take advantage of the available data
- We are living the information world? Or the raw data world?

How to collect mobility data?

- Laboratory setting
- Real-world setting
- Self-reported data collection method
- Using monitoring devices, sensors and accelerometers or using Internet of Things (IoT) devices

Laboratory-based Gait Monitoring

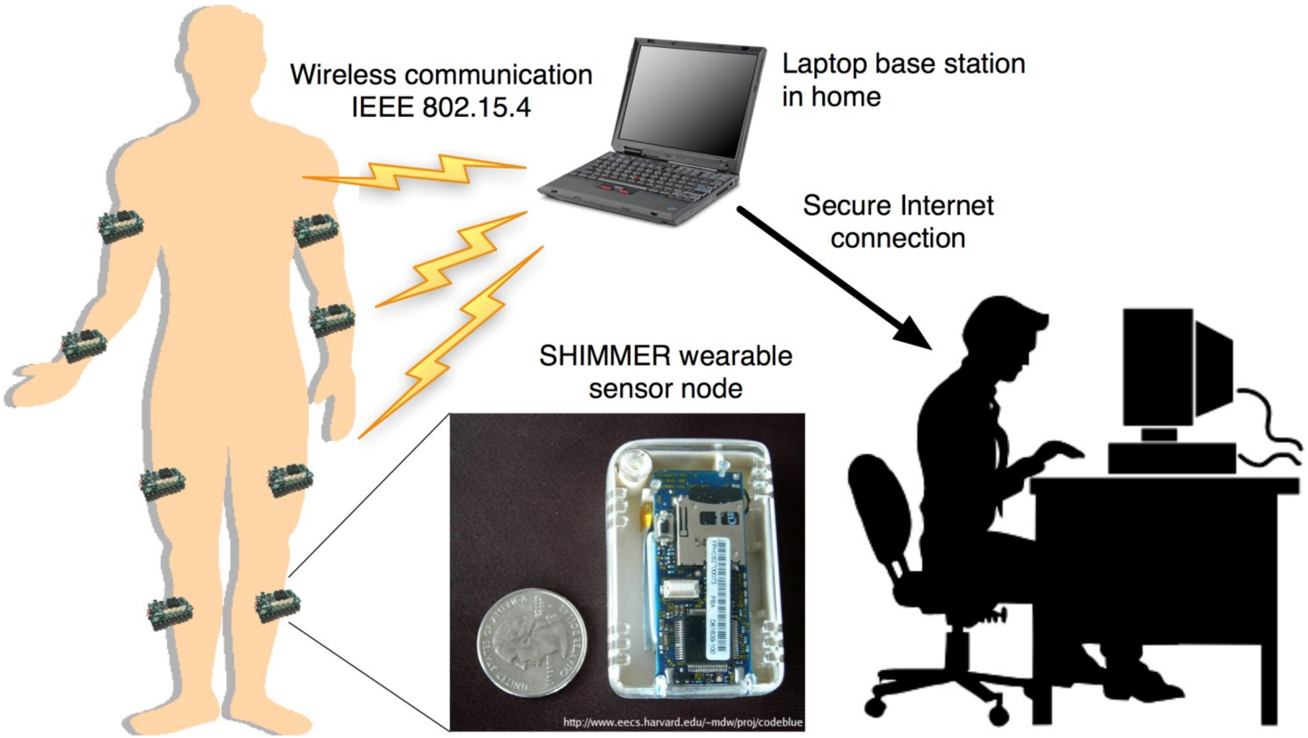
- Expensive
- Uncomfortable

- Limited mobility
- Complicated

Wireless Sensors and Mobility Monitoring

- Inexpensive
- Comfortable

- High mobility
- Simple



http://fiji.eecs.harvard.edu/Mercury

Challenges and Opportunities

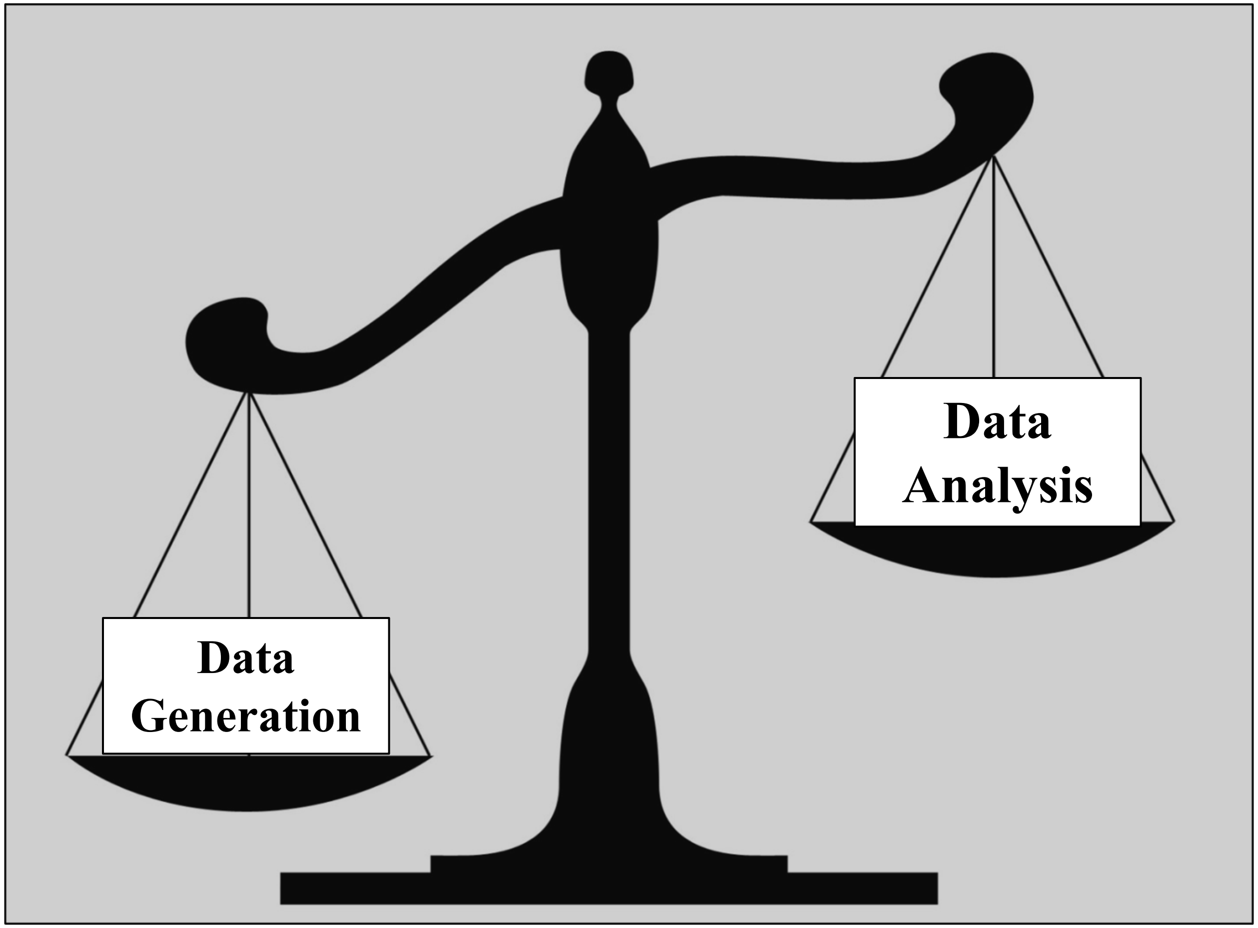
UNO BIOINFORMATICS

- Too much data?
- Are all available datasets relevant?
- Are they all accurate?
- Are they complete?
- Policy issues
- How to analyze such data
- Correlation versus causation
- How results can be verified? Validated?

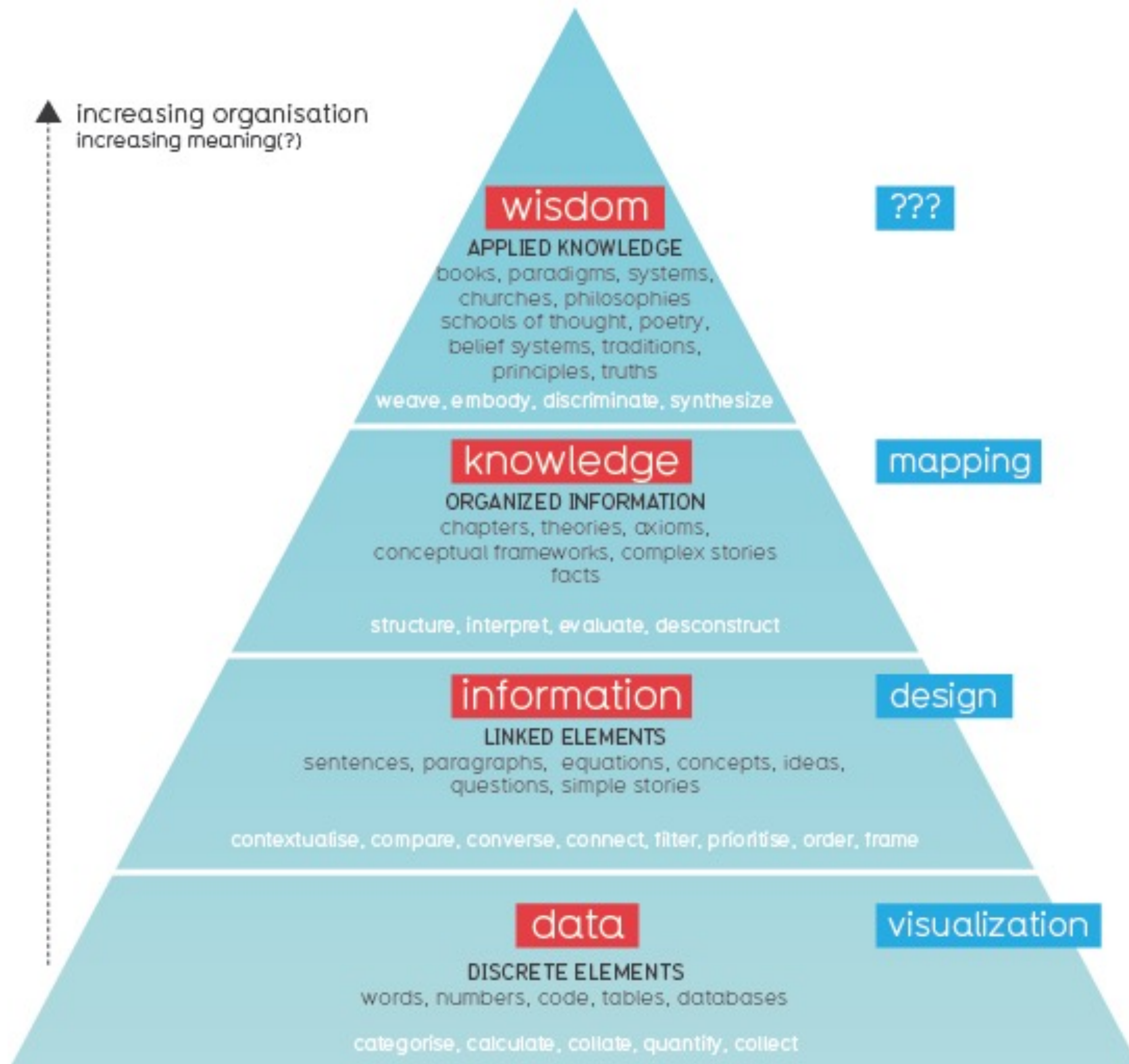
Data Generation vs. Data Analysis/Integration

- New technologies lead to new data:
  - Competition to have the latest technology
  - Focus on storage needs to store yet more data
- Biomedical community needs to move from a total focus on data generation to a blended focus of measured data generation (to take advantage of new technologies) and data analysis/interpretation/visualization
- How do we leverage data? Integratable? Scalable?
- From Data to Information to Knowledge to Decision making

Current Focus on Data Generation



Data-Information-Knowledge-Wisdom



http://www.ritholtz.com/blog/wp-content/uploads/2010/11/data_info_knowledge_wisdom.png

Smart Data Data-Driven Decisions

- Data: Physical entities at lowest abstraction level; contain little/no meaning – Measured data
- Information: Derived from data via interpretation – Processed data
- Knowledge: Obtained by inductive reasoning, typically through automated analysis and iterative collaboration – data + relationships
- Decision Support:

Big Biomedical Data

UNO BIOINFORMATICS

Health Data

- Patient history
- Allergies
- Prescription history
- Family history
- Surgeries

Social Data?

- Relationship/friendship data
- Location
- Smog
- Light Exposure

Genetic Data

- Personalized genome
- Genome *over time?*
- Susceptibilities
- Preventative therapeutics

23andMe

Wellness Data

- Sleep habits
- Eating habits
- Daily activity
- Stress levels

Mobility and Wellness

- Relationship between mobility and health is now well-documented
- Recent Advancement of Sensor technology Widespread of sensors
- Impact of Internet of Things
- So much relevant data is currently available
- Individual-focused mobility data
- The impact of the commercial aspect
- The potential of Big Data Analytics

Mobility and Health

- Mobility data collection: Continuous and non-invasive
- Mobility data: may not be 100% accurate
- Can we convert Individual-focused mobility data to population analysis
- From mobility parameters to health assessment to the prediction of health hazards
- Prevention and proactive healthcare as compared to reactive medicine

Widespread of Wearable devices

Main categories of wearable monitors:

- Pedometer
- Load transducer/foot-contact monitors
- accelerometers
- HR monitors
- Combined accelerometer and HR monitors
- Multiple sensor system

Motivation

- Can we use mobility parameters or patterns to assess health?
- Would mobility data be useful to adjust or tweak rehabilitation programs?
- Can we use mobility data to determine the right time to discharge patients after hospital procedures?
- Would mobility data be helpful in predicting health hazards before they happen?
- Can we utilize mobility information to determine the level of assistance needed for elderly in retirement facilities?

Applications in hospital: Post-operative Nursing Care

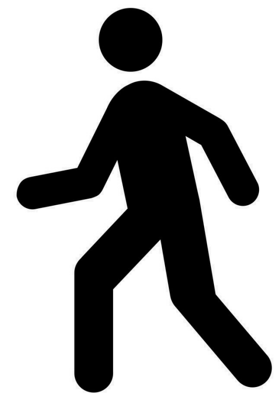
- A *post-operative* assessment is very important to a full and speedy *recovery from* any type of *surgery*.
  - a full assessment and an individualized treatment plan based upon the patient's needs and level of function, coupled with clinician expectations

Mobility Applications in Health

- The explosive and widespread use of sensors and wearable devices
- Many studies connected the way people move to their health (physical and mental), safety concerns and overall states of people and environments
- We can collect and store mobility levels parameters
- Data Analytics, AI and machine learning are critical in extracting knowledge from mobility data

Mobility & Health

- Inherent asset of every human being

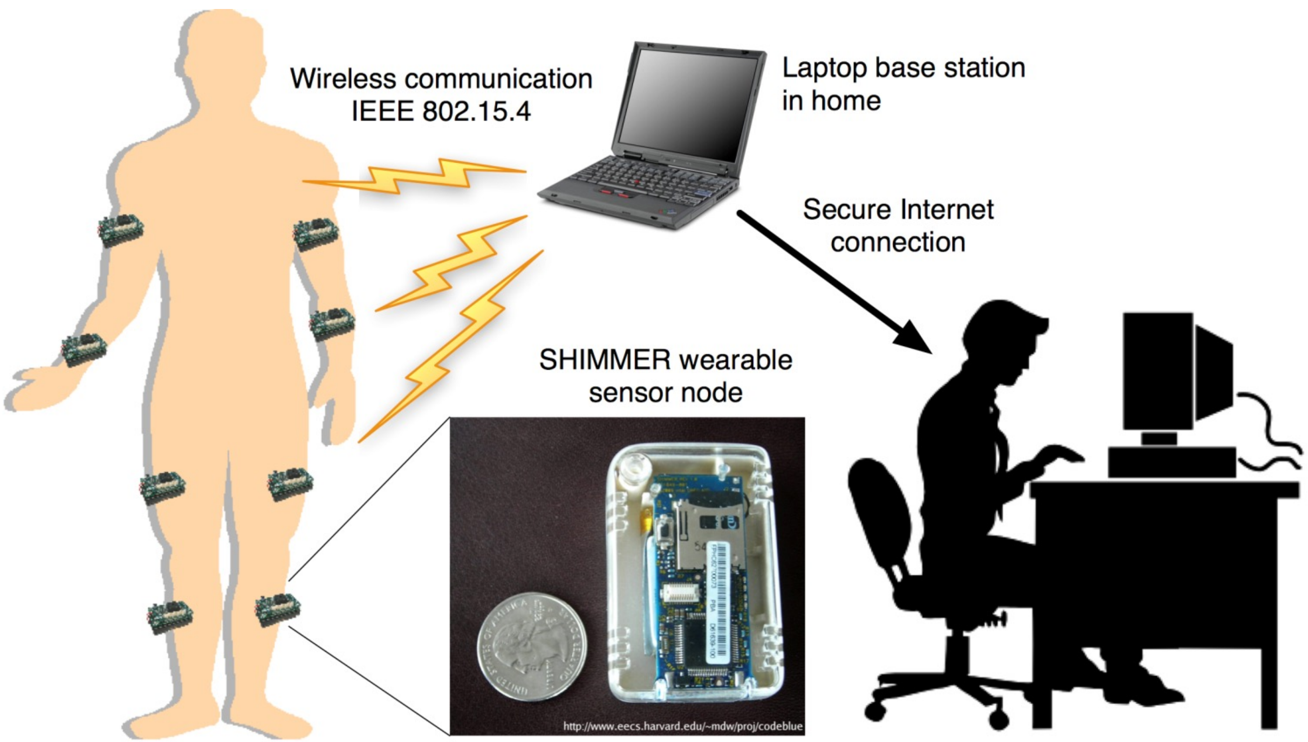


Wireless Sensors and network models

- Correlation between mobility and health level
- Monitoring mobility levels – can mobility prameters reflects health levels
- Aging of cells and aging of systems
- Collaboration between Bioinformatics researchers, Wireless Networks group and Decision Support Systems lab

Wireless Sensors and Mobility Monitoring

- Inexpensive
- Comfortable
- High mobility
- Simple



http://fiji.eecs.harvard.edu/Mercury

Mobility Monitoring

- Availability of many large useful devices – focus on collecting relevant data
- Availability of numerous helpful software packages
- Lack of data integration and trendiness of the discipline
- Fragmented efforts by computational scientists and domain experts
- Lack of translational work – from the research domain to engineering and healthcare applications
- Increasing interest among researchers, industry and educators

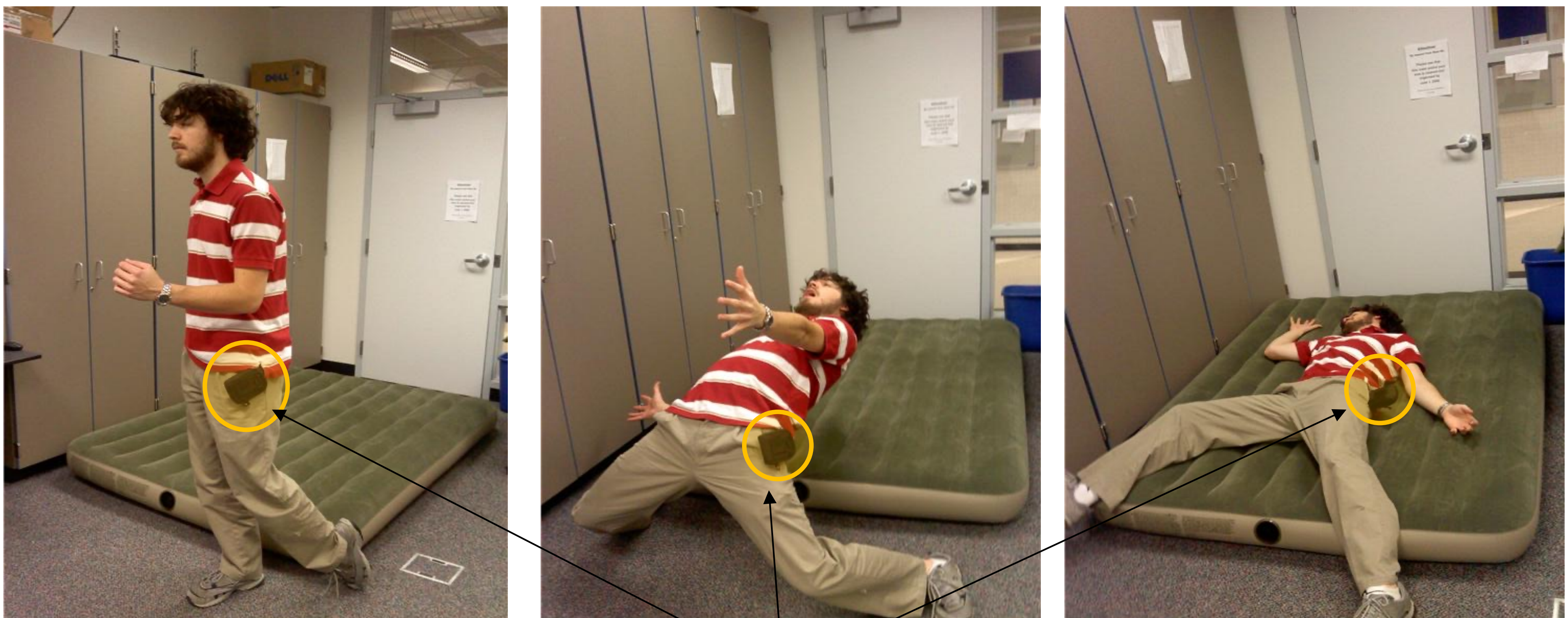
Goals of the Project

- Mobility Profile
  - Patient wearing a 3D-accelerometer will be monitored 24/7.
  - A complete mobility profile will be available for patients and care providers.
- Health hazards Prediction using Mobility Profiles
  - The system will identify anomalous movement and patterns that usually result in a fall or injury,
  - We would be able to take preemptive measures when such a pattern is detected, in order to reduce the occurrence of falls and prevent fall-related injuries.
  - We will develop an index that enables health care providers to determine how likely people are to fall.

Fall Detection

Accelerometer-based fall detection

- Measure acceleration in three orthogonal directions.



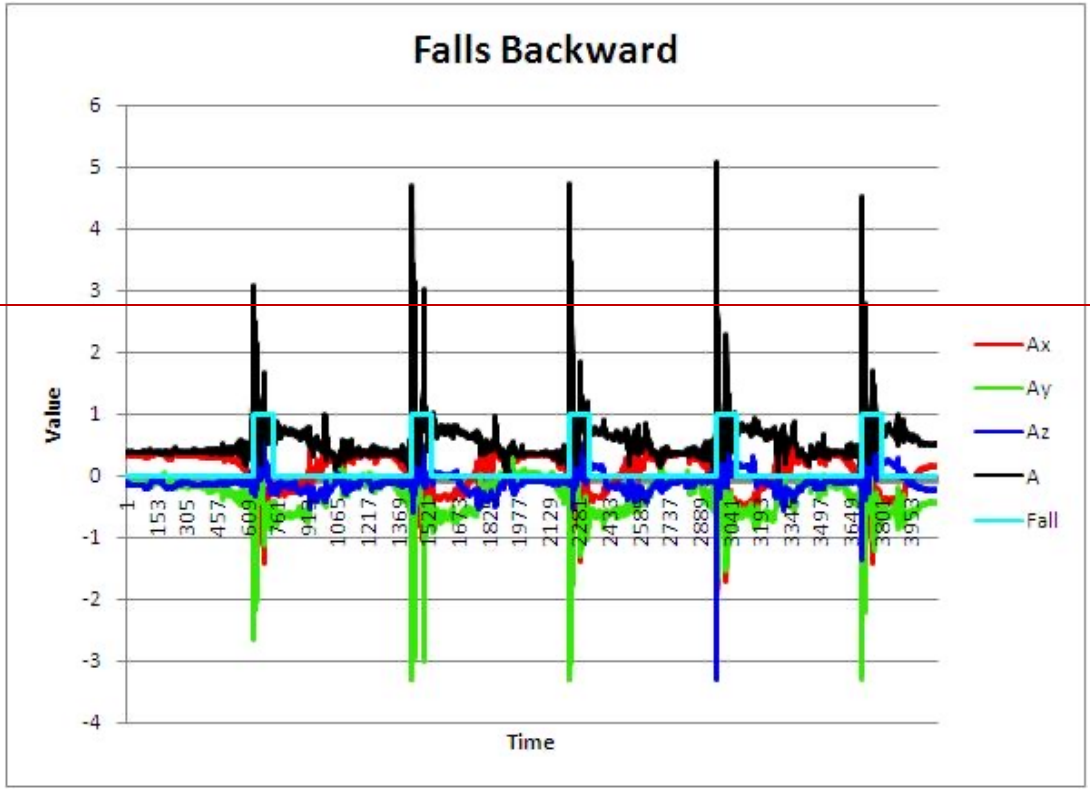
Shimmer Device

Fall Detection

Accelerometer-based fall detection

- Determine an acceleration threshold.
- Detect fall.

Threshold



How to Compare Mobility Patterns -Movement Words Coding Scheme

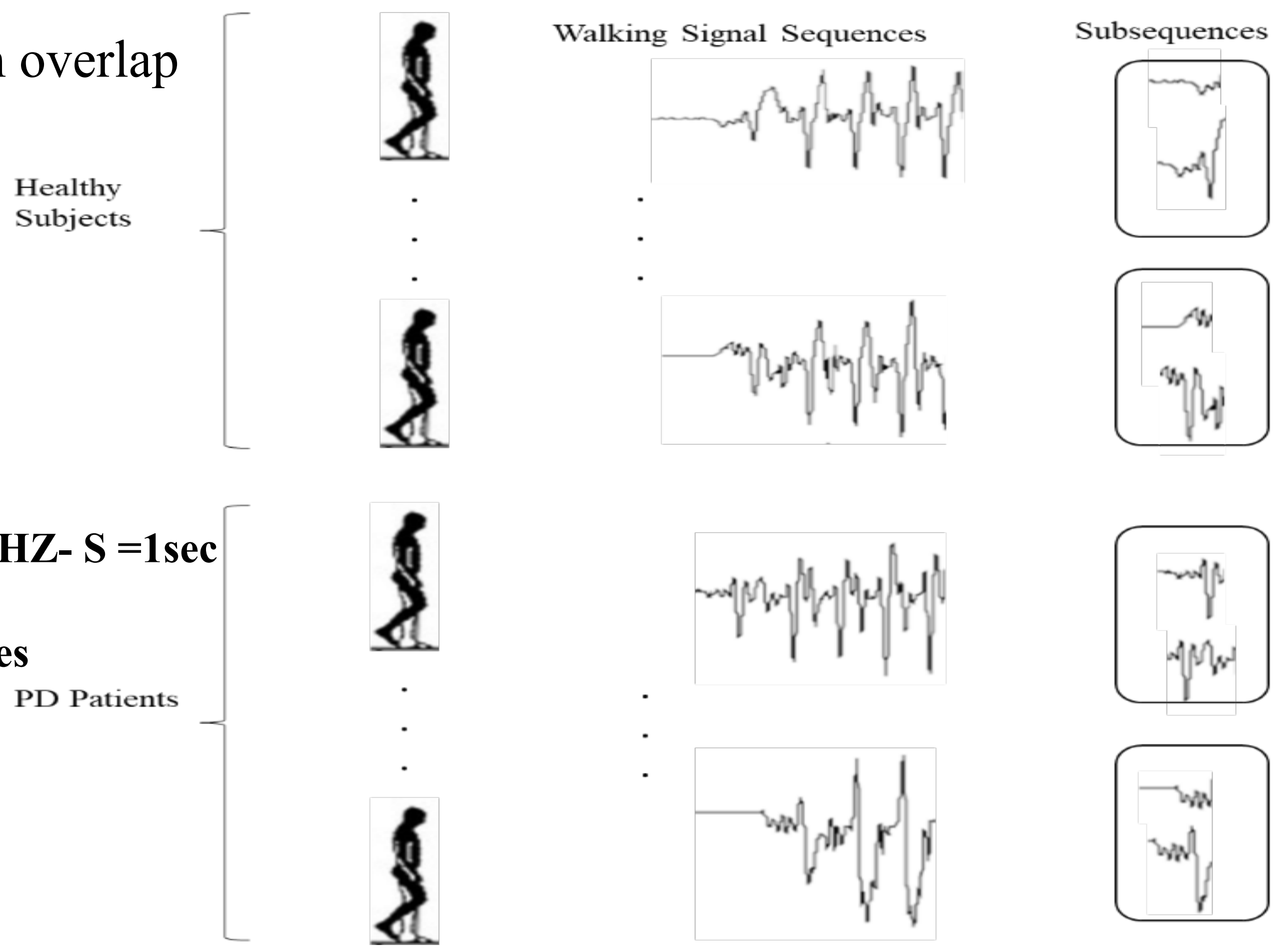
Generating Subsequences of Signal

Window Size = S sec with overlap

$$N = \sum_{i=1}^{n} 2 * \left(\frac{ti}{s}\right) - 1$$

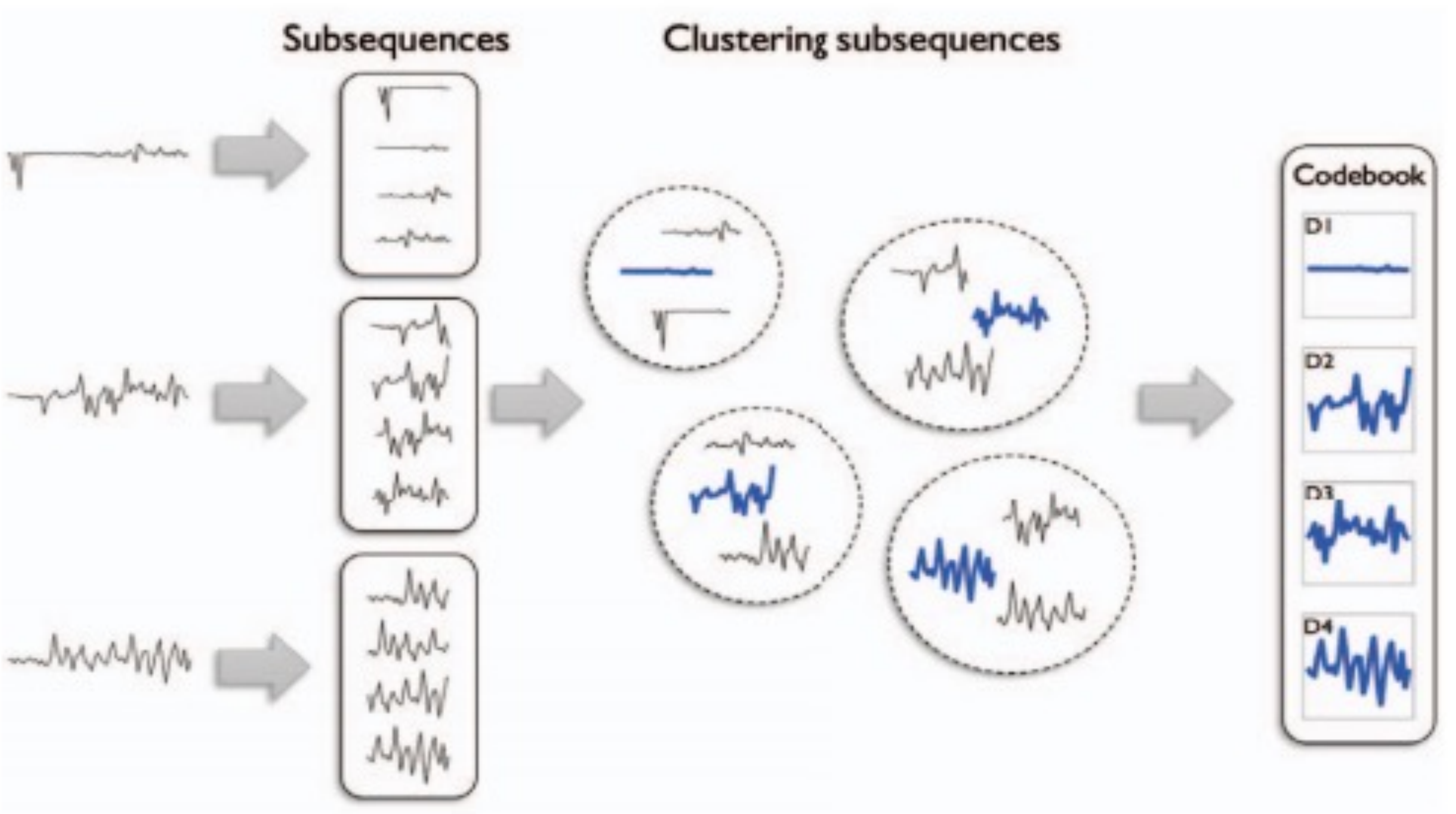
1 Person-1 Day- Frequency 100HZ- S =1sec

8,640,000 subsequences

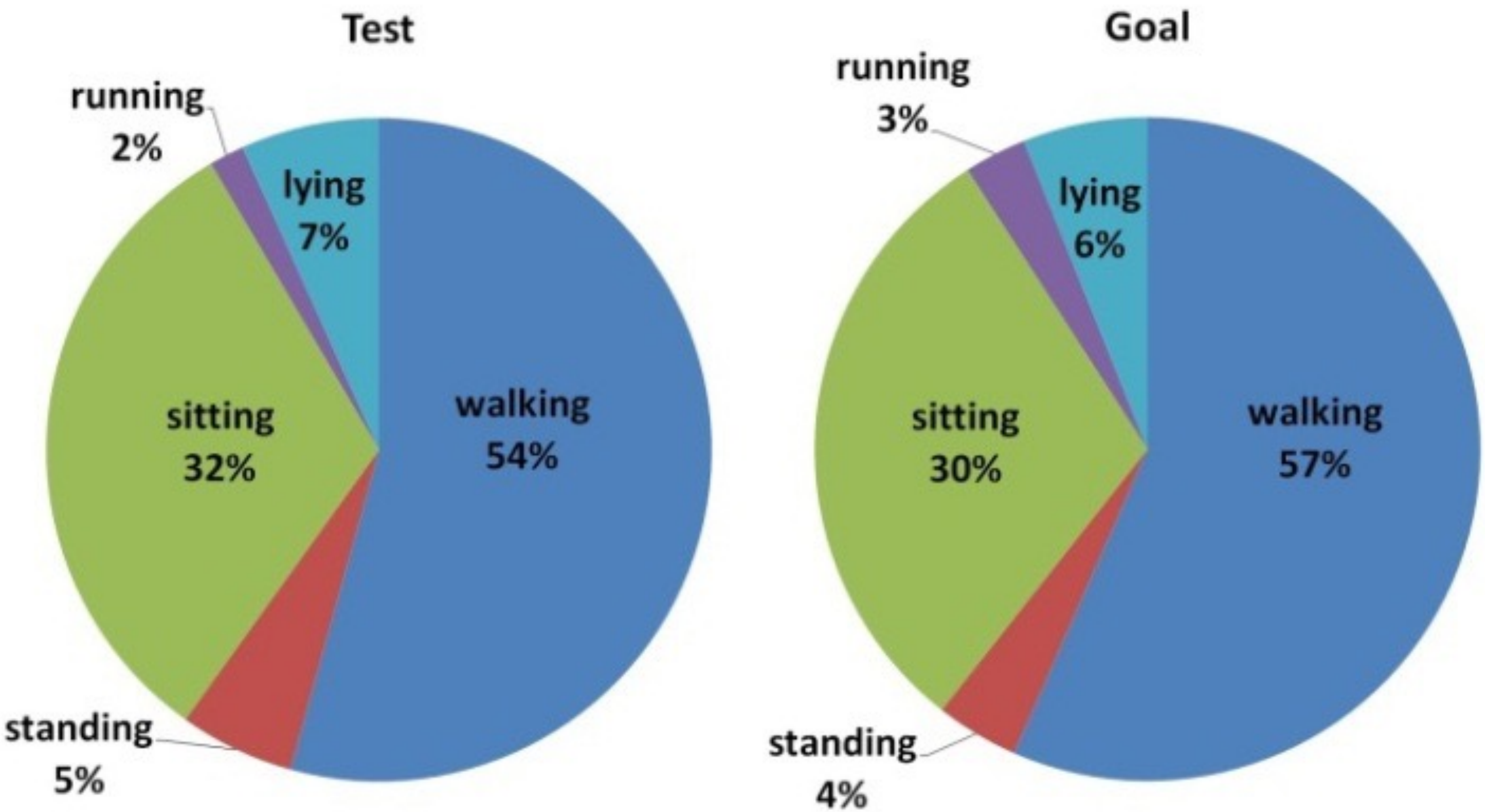


Feature Engineering- Movement Words Coding Scheme

Vocabulary Generation

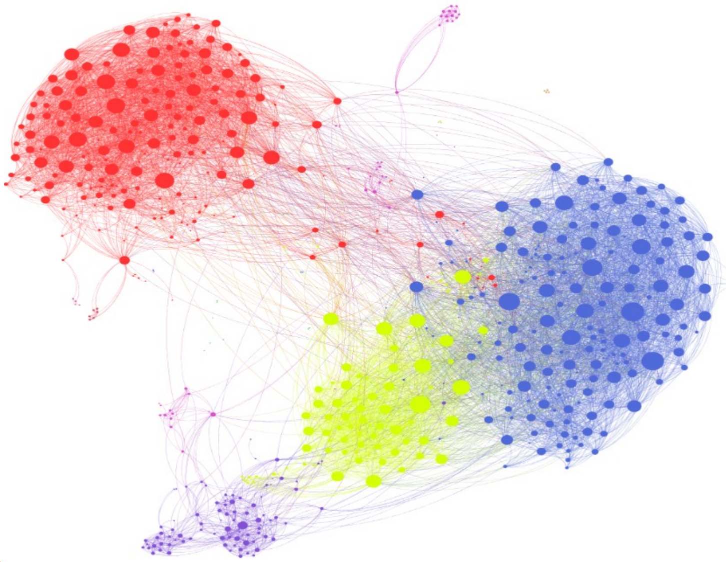


Mobility Profiles



Target Populations

- Parkinson's Disease
- Multiple Sclerosis
- Amyotrophic Lateral Sclerosis
- Huntington's disease
- Aging (Geriatrics)
- …
- Healthy Control (age-matched)



Dataset: Participants and Protocol (Ankle Data)

- Protocol:
  - 4 minute Walking (around the hospital)
  - Sampling frequency:100
  - **Moderate PD**

	Control	**PD**	**Geriatrics**
Number of subjects	5	5	5
Gender (M/F)	3:2	3:2	2:3
Age	64 ± 10	72 ± 6.3	81 ± 5.9
UPDRS III		20.8 ± 6.1	
H & Y		2.6 ± 0.5	

Dataset: Participants and Protocol (Wrist Data)

- Three phases of data collection (6-months period between each two phases)- One week of data per individual-per week
- Sampling frequency:100
- Mild, moderate, and sever PD (overall mild PD)

	Healthy young	Healthy elderlies	PD
Number of subjects	3	3	3
Gender (M/F)	2:1	1:2	1:2
Age	23 ± 3.6	65.3 ± 16.2	66 ± 5.0
UPDRS III		------	
H & Y		2.16 ± 0.88	

Dataset: Participants and Protocol (Wrist Data) Second Phase

- Sampling frequency:100
- Mild, moderate, and sever PD (overall mild PD)

	Healthy young	**Healthy elderlies**	**PD**
Number of subjects	19	23	17
Gender (M/F)	12/7	10/13	14/3
Age	24.5 ± 2.1	63.6 ± 7	72.1 ± 6.4
UPDRS III		------	
H & Y			1.29 ± 0.5

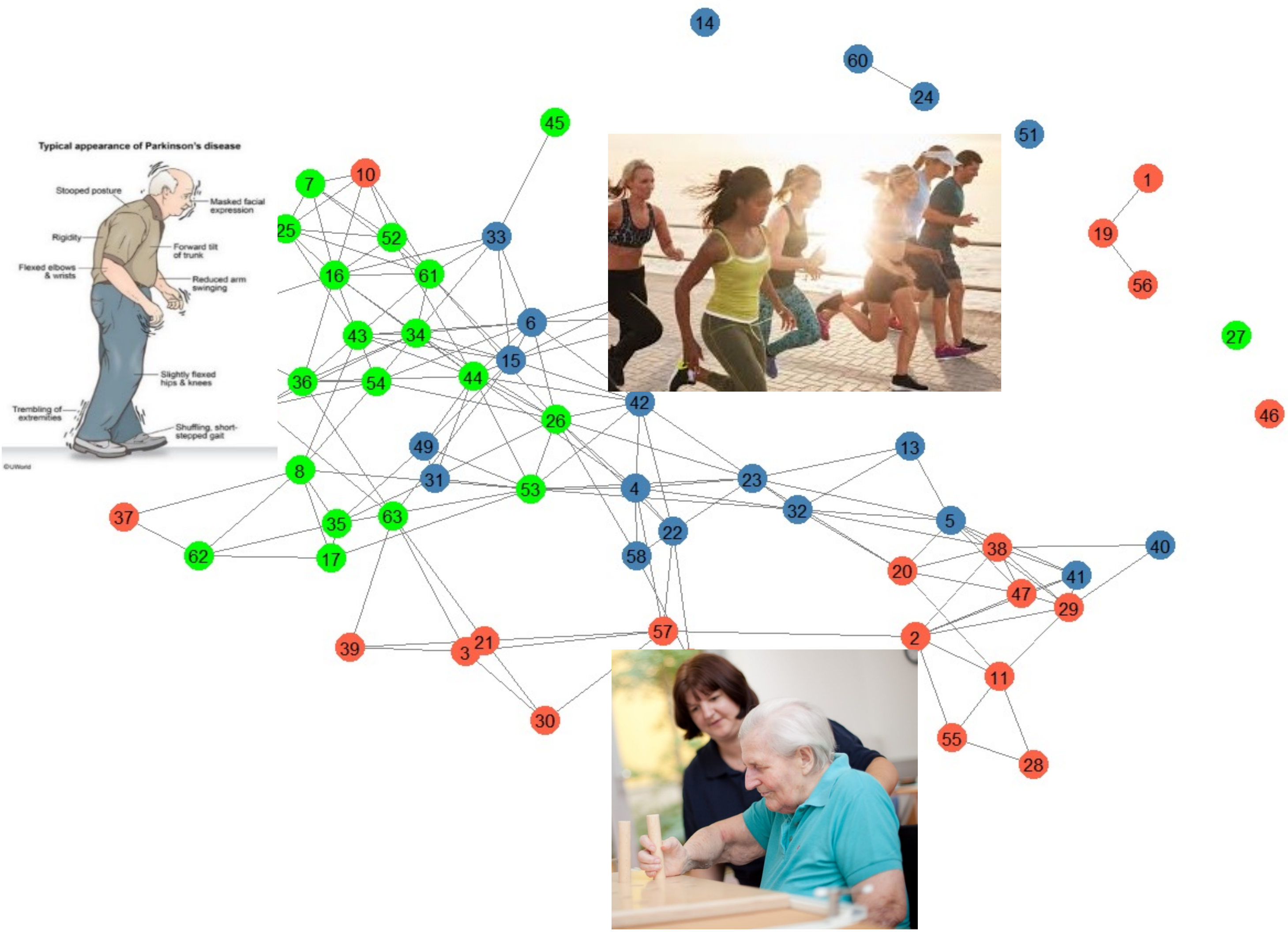
Selected Set of Features

Feature's Name	Description	Feature category
Variabiltiy_StrideTime	Variability of stride time	Signal level
Variability_SVM	Variability of vector magnitude	Signal level
Variability_RMSX	Variability of root mean square in the AP direction	Signal level
Variability_RMSZ	Variability of root mean square in the ML direction	Signal level
Velocity	velocity	Signal level
Smoothness_X	Smoothness in the AP direction	Signal level
Smoothness_Z	Smoothness in the ML direction	Signal level
RMSZR	Root mean square relative to the mean value in the ML direction	Stride level

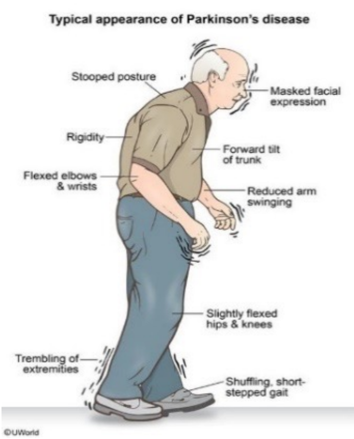
Modeling: Machine Learning

- **Standard Features:**
  - All features (32)
  - First reduced set of features (22)
    - Using Information Gain and Ranker methods
  - Second reduced set of features (8)
    - Using Pearson Correlation coefficient and ANOVA table
  - Third reduced set of features (7)
    - feature sets with one feature less than the optimal number of features
- **Document-of-Words Features:**
  - 10 Features for wrist data and 4 features for ankle data
- **Various Machine Learning Techniques:**
  - SVM, Random Forest, Naïve Bayes, AdaBoost, and bagging
- **Validation:**
  - K-Fold Cross validation
- **Accuracy measures:**
  - F-measure, Precision, Recall

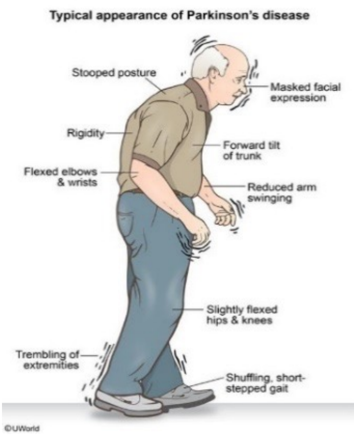
Similarity Network Model – Wrist Data- Word Features

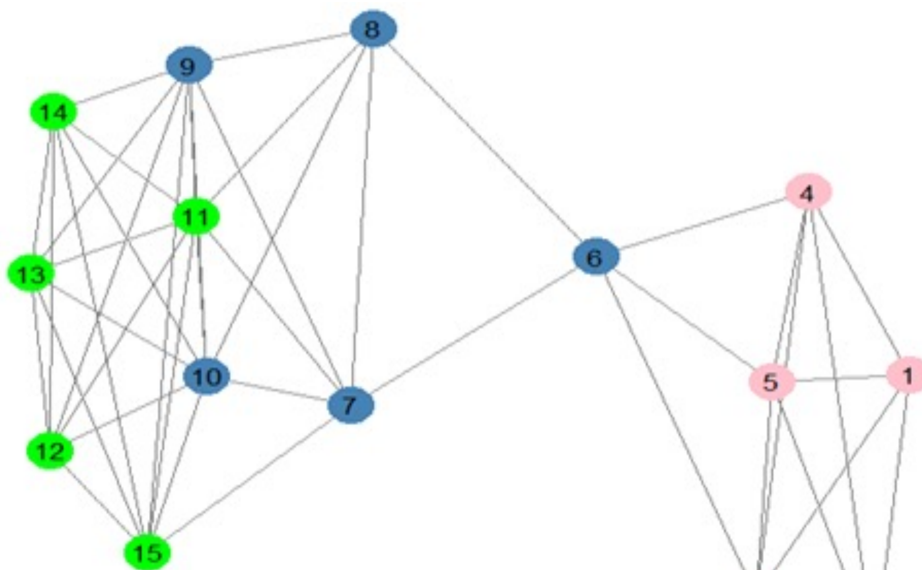


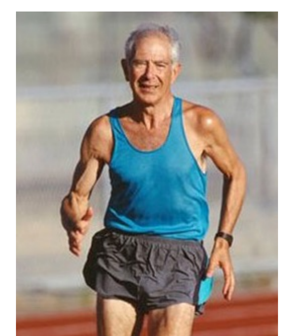
Similarity Network Model- Ankle Data (Mild PD) All Features



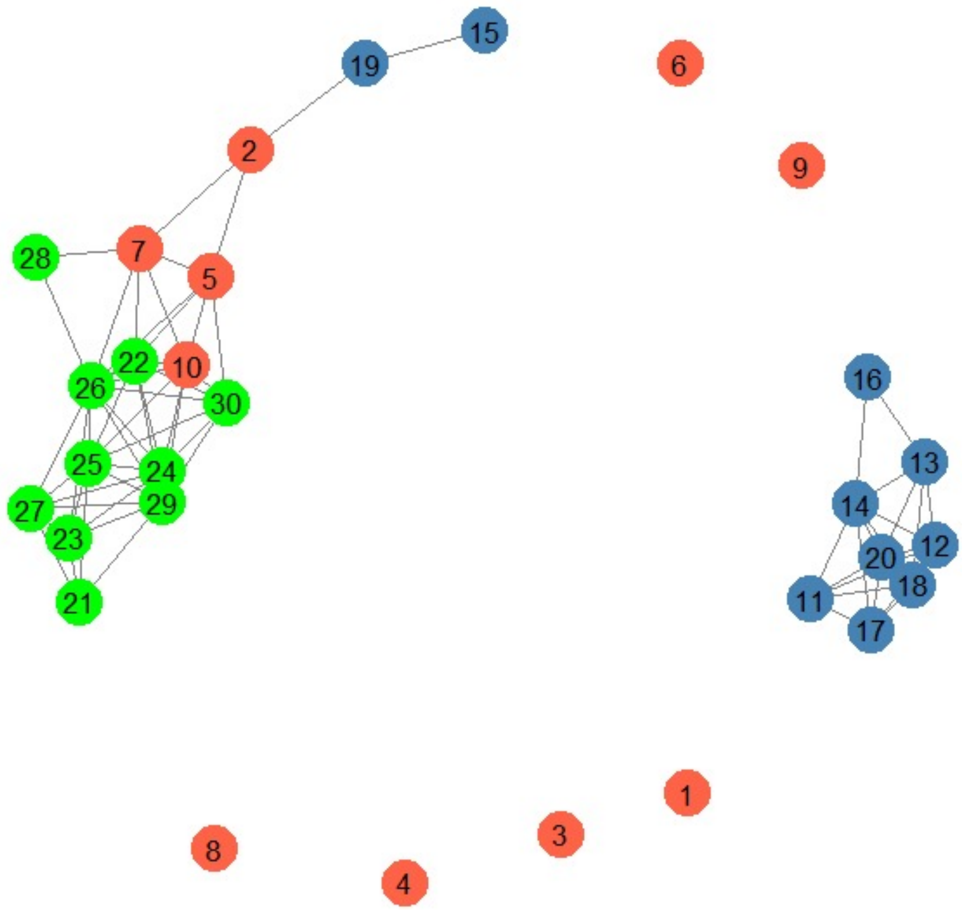
Similarity Network Model- Ankle Data (Moderate PD)-All Features



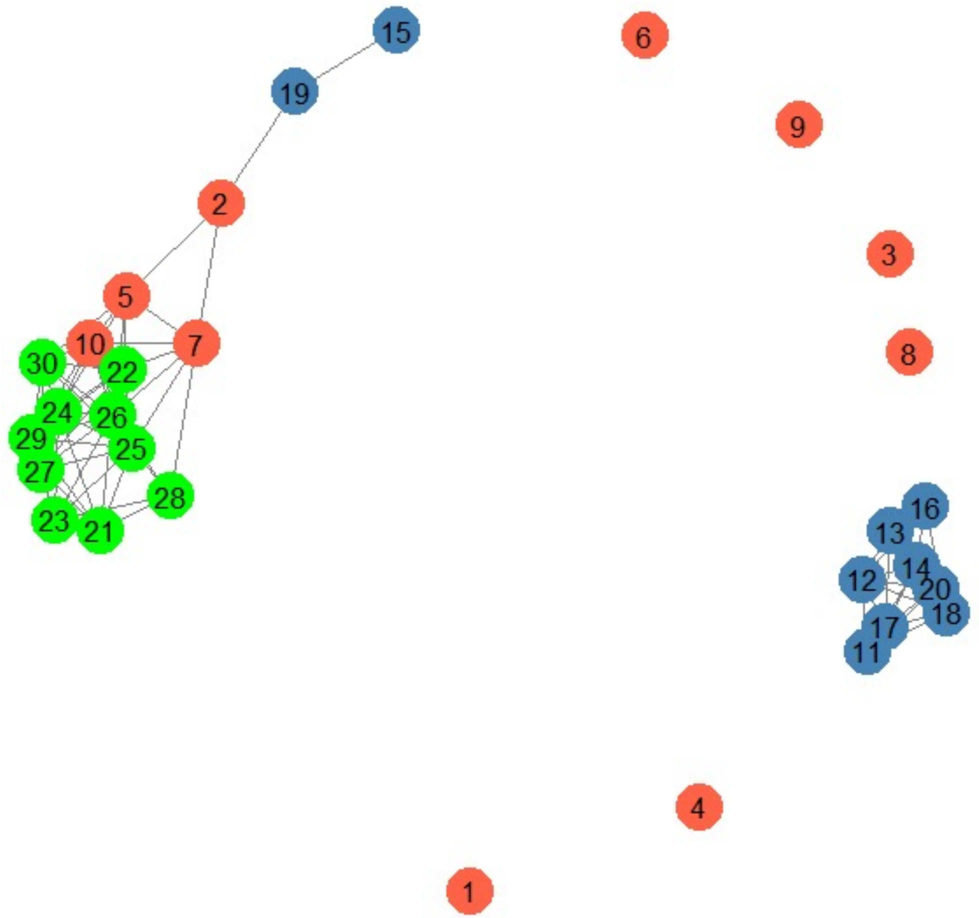




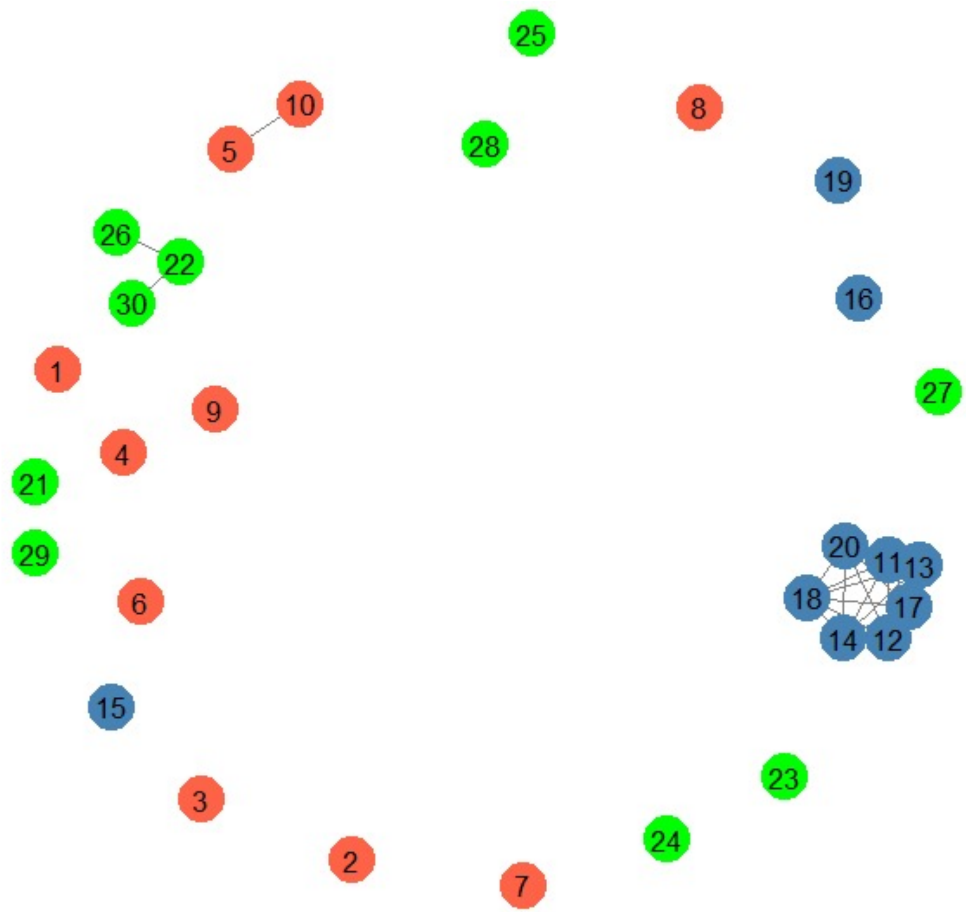
Similarity Network Model-Ankle Data (Mild PD)- Reduced_22



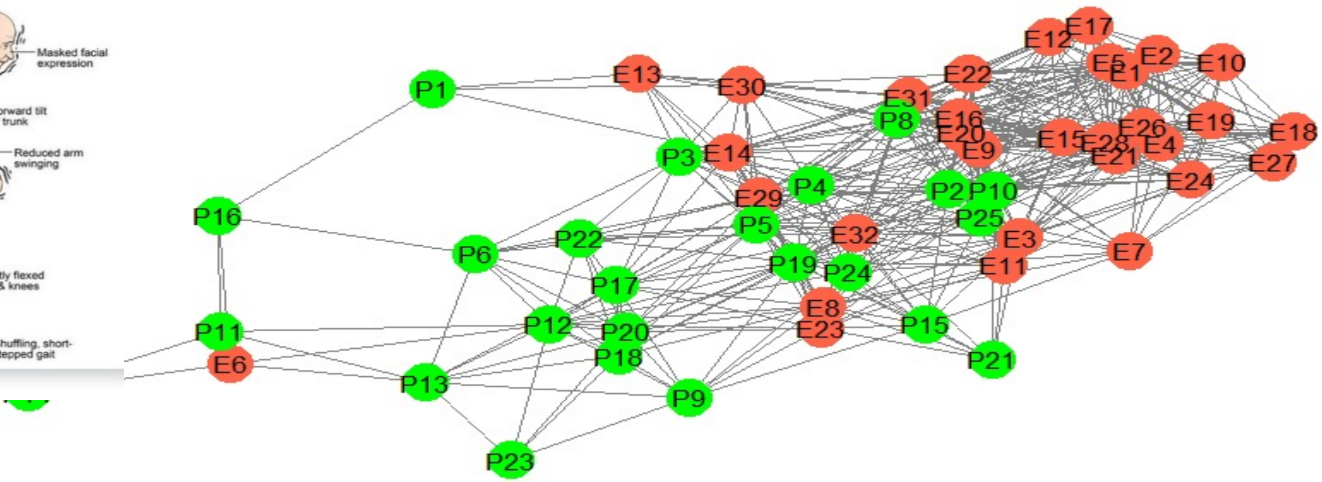
Similarity Network Model-Ankle Data (Mild PD)- Reduced_8



Similarity Network Model-Ankle Data (Mild PD)- Reduced_7



Similarity Network Model for the data from the first phase of wrist dataset- Threshold at **90%-** PD and HE



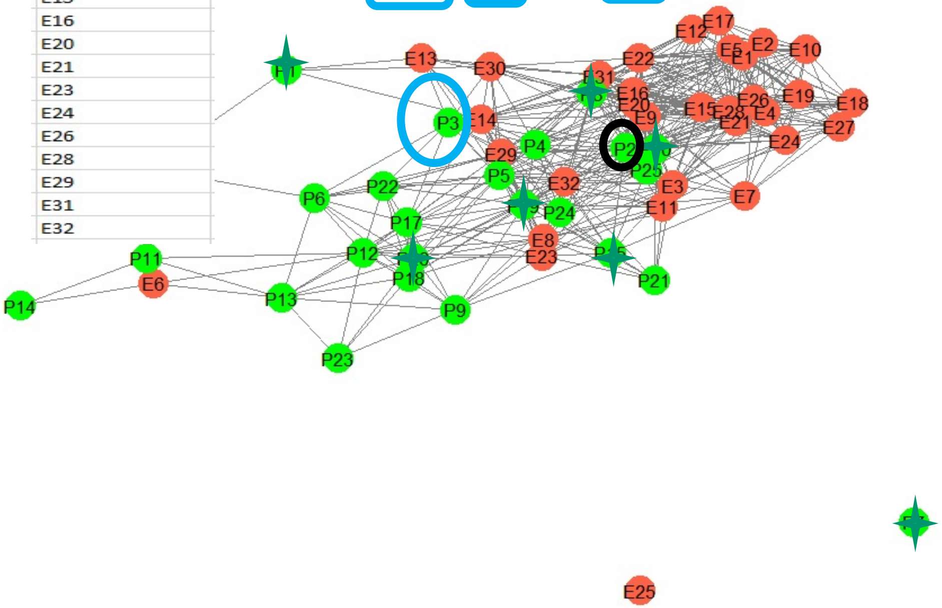
Subject	Gender	Age	MoCA	FoG	FAB	TUG	GDS	H&Y	MFES	Lawton
PD8	Male	69	28	2	39	6.7	0	1	10	8
PD10	Male	71	28	1	39	6.7	1	1	9.3	8
PD1	Male	83	26	8	39	11.2	0	1	8.6	8
PD21	Male	54	25	0	39	9.0	0	1	10	8

Phase1 ~ Phase2- P2 & P3 (90%)

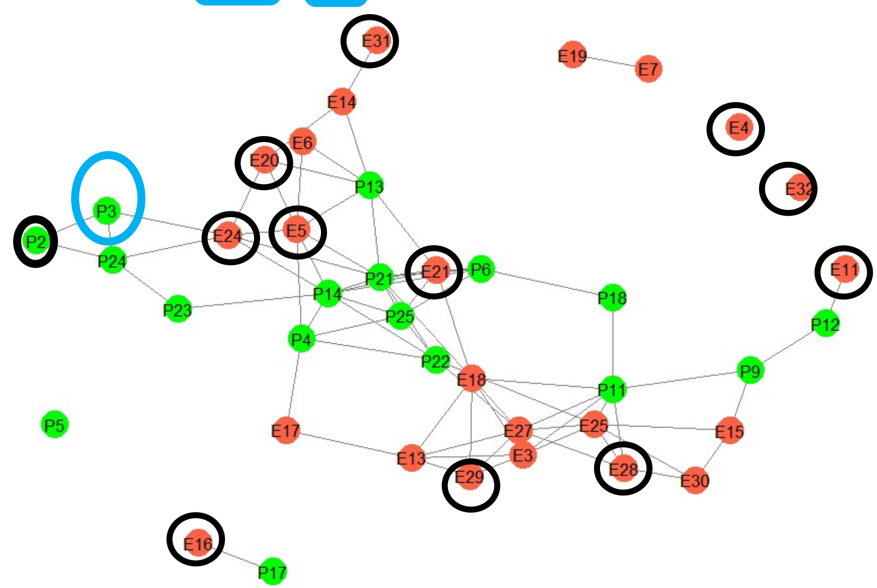
UNO BIOINFORMATICS

Connected to P2
E1
E2
E3
E4
E5
E8
E9
E1
E15
E16
E20
E21
E23
E24
E26
E28
E29
E31
E32

Gender	Age	Falls in	MOCA S	FOG !	FAB sc	TUG T	Geria	H-Y	MFES /	Lawto
Male	83	0	26	8	39	11.1	0	1	8.57	8
Male	68	0	29	1	37	7.25	1	1	10	8
Male	69	0	19	3	37	7.99	2	1	9.86	8



Gender	Age	Falls i	MOCA	FOG Sc	FAB scoi	TUG Tin	Geriac	H-Y sc	MFES Av	Lawto
Male	68	0	22	4	39	8.64	0	1	10	8
Male	70	0	27	4	38	10.225	2	1	9.64	8



Post-operative Nursing Care

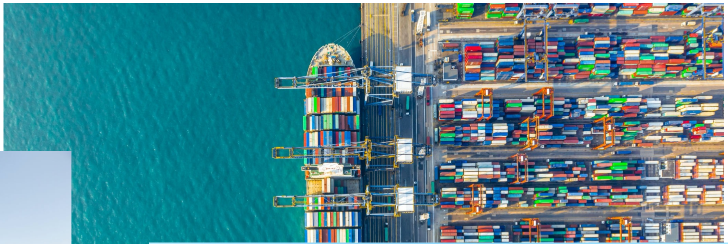
- A *post-operative* assessment is very important to a full and speedy *recovery from* any type of *surgery.*
  - a full assessment and an individualized treatment plan based upon the patient's needs and level of function, coupled with clinician expectations

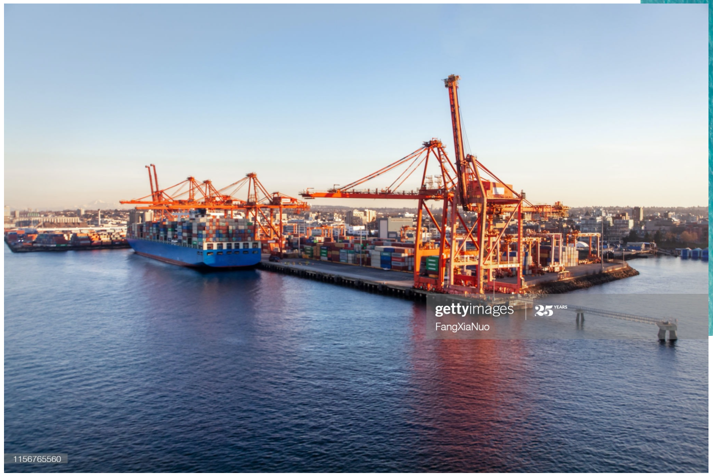
Applications for health subject: Physical therapy / Rehabilitation

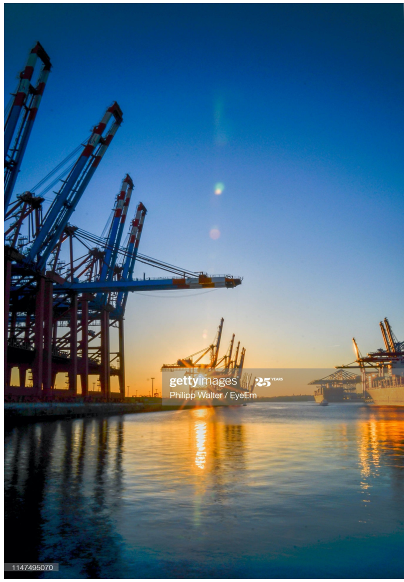
- Help a patient perform rehabilitation exercises to improve their balance and mobility, and
- Find exercises that meet patient's specific needs and abilities.

Mobility in Ports and Marine Applications

- Efficient movement in moving platforms represents additional challenges - proper assessment is completely missing in such applications
- Many critical functions in such environments depend on mobility parameters.
- Network models can be used to assess workers' ability to adjust to moving grounds.

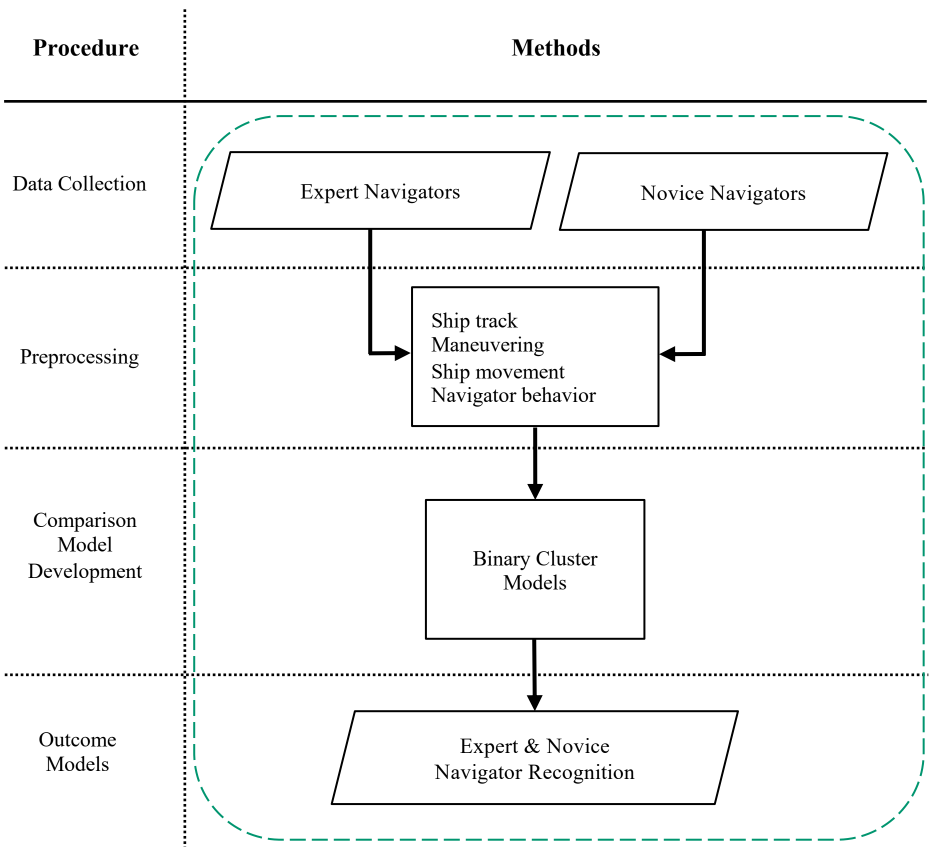


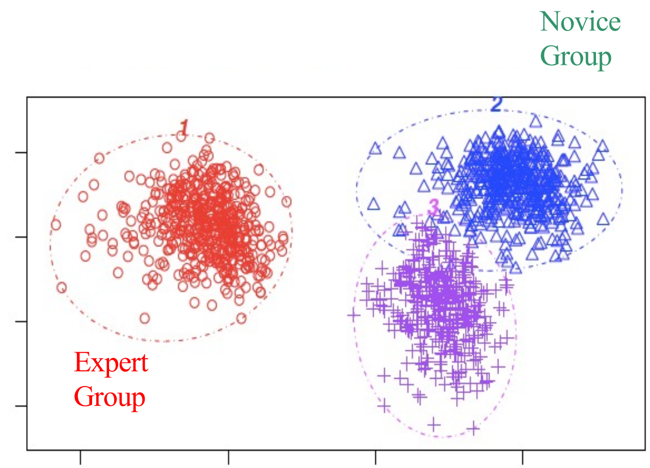


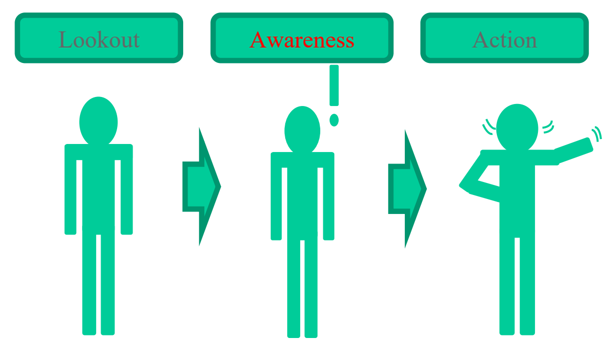


Expert & Novice Navigator Recognition

Explore the differences human behavior base on proficiencies





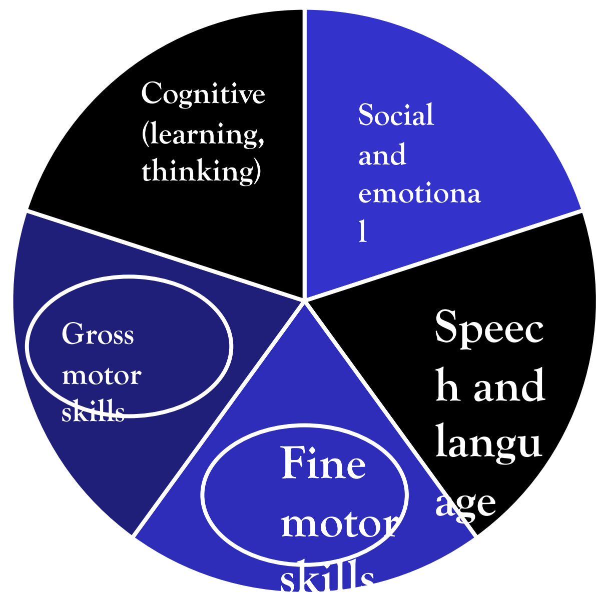


Summary of Case Study

- Correlation Network model worked beautifully when we applied it to both dataset.
- MIGMC provided us with the best set of features
- Accelerometers at ankles and Wrist can capture gait parameters that are useful in early diagnosis of disease.
- The performance of Bag-of-Words model is ,if not higher than, equal to the Standard Model
- Ankle data are more precise in identification of patients with PD compared to Wrist Data
- Still wrist could be argued as a better body location (87.5% accuracy is good enough)

Mobility Developmental Disorder

- Developmental disorder during early childhood influences motor/mobility
  - Autism
  - Cerebral Palsy
- Certain diseases at any stage in the lifetime may alter mobility
  - Mental disorders - Depression



5 types of skills developed during early childhood

Existing clinical diagnosis

- Developmental disorders (Autism, CP)
- Mental disorders (Depression)
- No pathology test
- Self reporting assessment
- Interview based diagnosis
- Observational scale
  - MADRS score for Depression
  - Pretchl's assessment for CP

Existing clinical diagnosis

- Existing approaches are time intensive
- Accuracy is observer dependent
- Requires expensive clinical infrastructure
- Helpful tool for clinicians for accurate decision making
- Necessity of sophisticated approach

Mobility as a feature

- Motor activity of conditioned is lower than their healthy counter parts
- Mobility has been used to characterize patients with movement disorders
- Can we also characterize or assess developmental disorders by utilizing mobility?

Analysis of Depression Episodes

- Depression
  - 280 million people suffering from depression across the world[1]
  - a deep sorrow that lasts for more than two weeks
  - Serious mental disorder
  - Poor performance in work, school
  - Affects quality of life, relationships

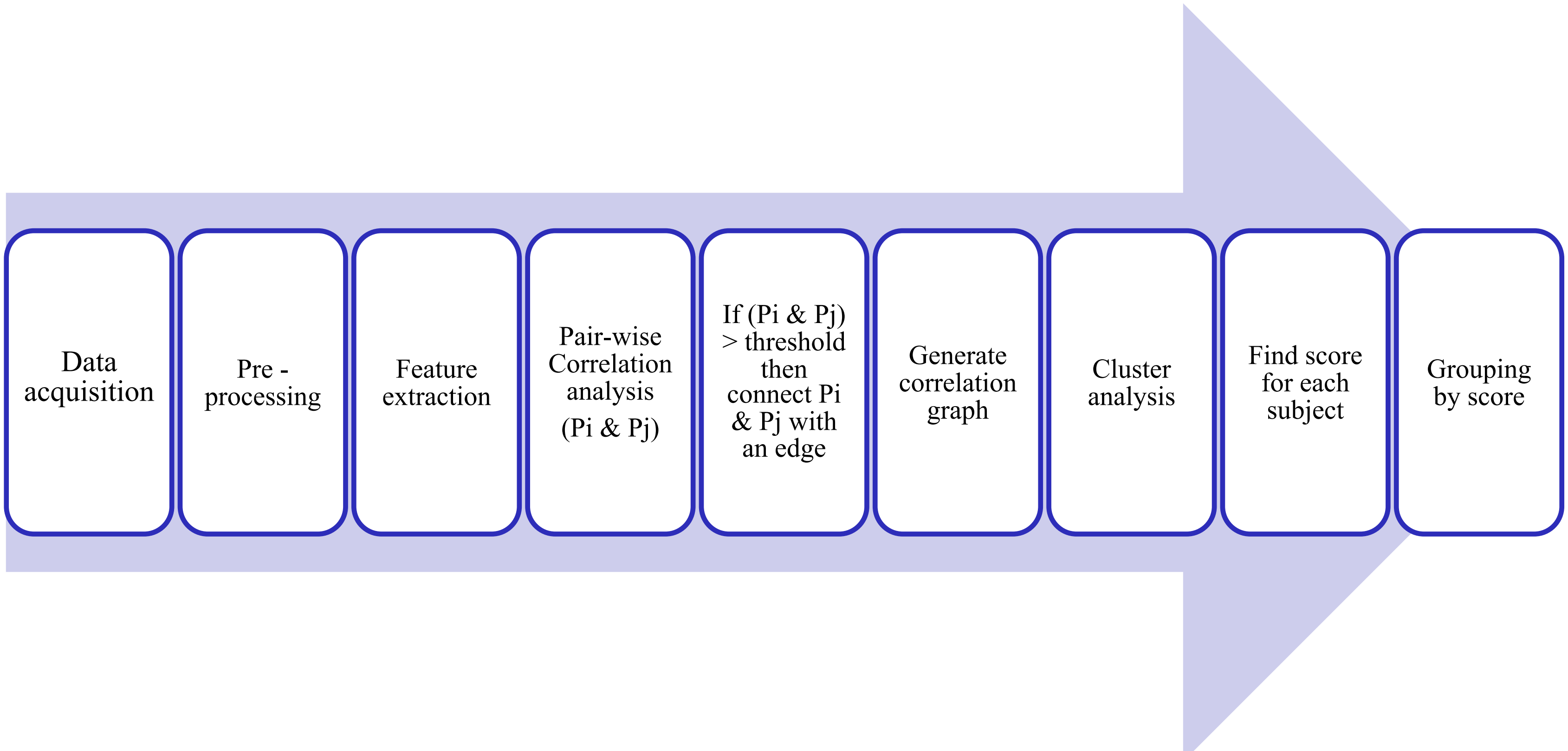
World Health Organization https://www.who.int/news-room/fact-sheets/detail/depression

State-of-the-art Clinical diagnosis

- No standard pathology test
- Feedback based assessment
  - diagnostic questionnaire assessed by clinical expert
  - Score based tests
  - Score indicates severity of the symptoms
  - E.g. Montgomery-Asberg Depression Rating Scale (MADRS)

. Montgomery, S.A. and Asberg, M., A new depression scale designed to be sensitive to change, Br. J. Psychiatry, 1979, vol. 134, no. 4, pp. 382–389.

Correlation Network Model



Depression Public Dataset

- Depression dataset
- 55 Subjects- 23 condition (depressed) & 32 control (normal)
- Duration – avg 12.6 days by each subject
- Sensor – ActiWatch accelerometer-based sensor worn on right wrist
- Data - each person activity is recorded in separate csv file with time stamp

3. Enrique et all, Depression: A Motor Activity Database of Depression Episodes in Unipolar and Bipolar Patients, In MMSys'18 Proceedings of the 9th ACM on Multimedia Systems Conference, Amsterdam, The Netherlands, June 12 - 15, 2018.

Depression public dataset

- Sensor records integration of intensity, amount and duration of movement in all directions (in the form of actigraph count)
- number of counts is proportional to intensity of the movement.

timestamp	date	activity
5/7/2003 12:00	5/7/2003	0
5/7/2003 12:01	5/7/2003	143
5/7/2003 12:02	5/7/2003	0
5/7/2003 12:03	5/7/2003	20
5/7/2003 12:04	5/7/2003	166
5/7/2003 12:05	5/7/2003	160
5/7/2003 12:06	5/7/2003	17
5/7/2003 12:07	5/7/2003	646
5/7/2003 12:08	5/7/2003	978
5/7/2003 12:09	5/7/2003	306

Demographic details

- Demographic details of each patient
  - Number of days monitored by each patient
  - Gender
  - Age
  - Unipolar or bipolar disorder
  - MADRS score at day 1 and day last

number	days	gender	age	afftype	melanch	inpatient	edu	marriage	work	madrs1	madrs2
condition_	11	2	35-39	2	2	2	10-Jun	1	2	19	19
condition_	18	2	40-44	1	2	2	10-Jun	2	2	24	11
condition_	13	1	45-49	2	2	2	10-Jun	2	2	24	25
condition_	13	2	25-29	2	2	2	15-Nov	1	1	20	16
condition_	13	2	50-54	2	2	2	15-Nov	2	2	26	26
condition_	7	1	35-39	2	2	2	10-Jun	1	2	18	15
condition_	11	1	20-24	1	NA	2	15-Nov	2	1	24	25
condition_	5	2	25-29	2	NA	2	15-Nov	1	2	20	16

Research questions

- Building a graph where groups of persons with similar mobility patterns are strongly connected
- Identify clusters/communities exhibiting similar mobility profiles
- Group according to their inter & intra cluster properties
- Enrichment analysis

Features

TABLE II: Hour-wise features

Feature name	**Features count**	**Feature description**
m0-m23	24	Mean (average) of motor activity measured for every hour for 0-23 hours
sd0-sd23	24	The standard deviation of motor activity measured for every hour for 0-23 hours
id	1	The unique id represents an individual from 55 subjects

Constructing a correlation graph

- Pearson correlation coefficient is applied on 55x55 feature matrix for each model

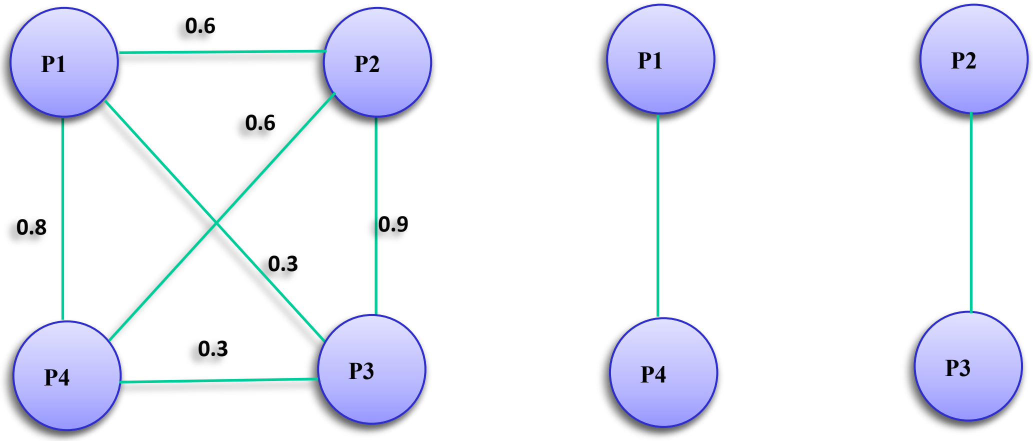


Figure : (a) Pairwise correlation (b) Correlation graph generation with threshold of 0.8

4. J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in Noise reduction in speech processing. Springer, 2009,pp. 1–4.

Clustering/ discovering communities

- Community is group of one or more persons exhibiting similar mobility characteristics [5]
- MCL(Markov Clustering) algorithm[6] with default parameters (expansion = 2 and inflation =2)
- MCL works by randomly visiting to find the strongly connected components in the graph.

5. M. Girvan and M. E. Newman, "Community structure in social and biological networks," Proceedings of the national academy of sciences,vol. 99, no. 12, pp. 7821–7826, 2002.
6. S. Dongen, "A cluster algorithm for graphs," CWI (Centre for Mathematics and Computer Science), 2000.

Cluster analysis and grouping

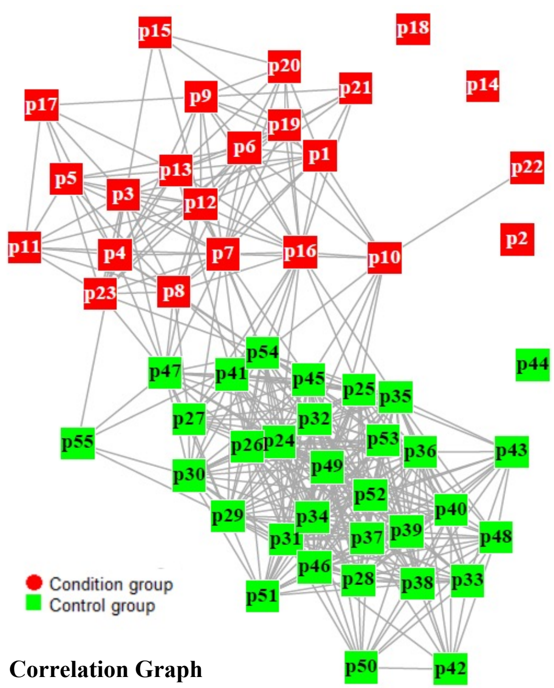
- Score (P_i) = (Σ inter cluster edges) / (Σ intra cluster edges)
  - For all subjects
- Score is representation
  - how well the node is connected to other nodes in the same cluster (Homogeneity)
  - How good the node is far from different cluster (Separability)

Cluster analysis and grouping

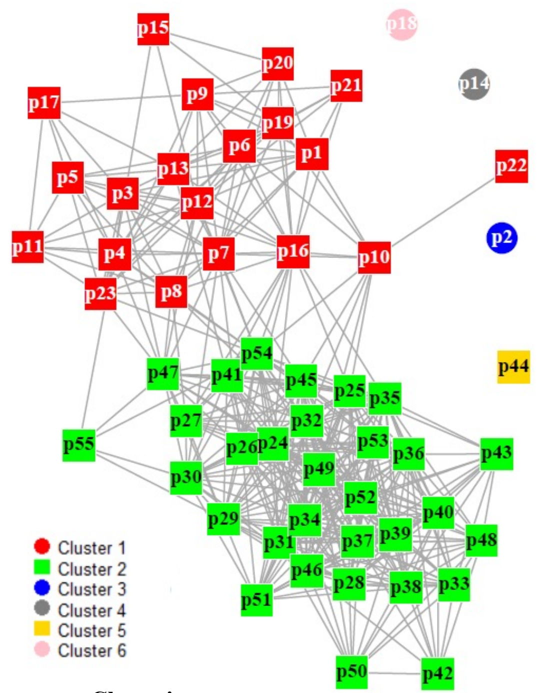
- Group the nodes in the same cluster into two groups

Group no	Original group	score	Interpretation
Group 1	Condition	0	Depression severity is higher
Group 2	Condition	>0	Depression severity is lower
Group 3	Control	>0	healthy but similarity is higher with respect to condition group
Group 4	Control	0	healthy group

Correlation graph and clustering

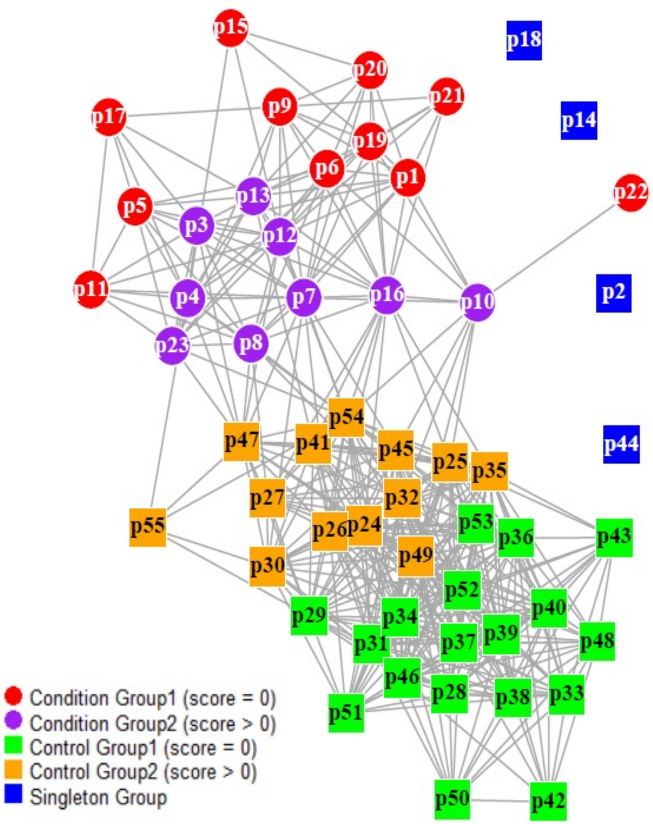


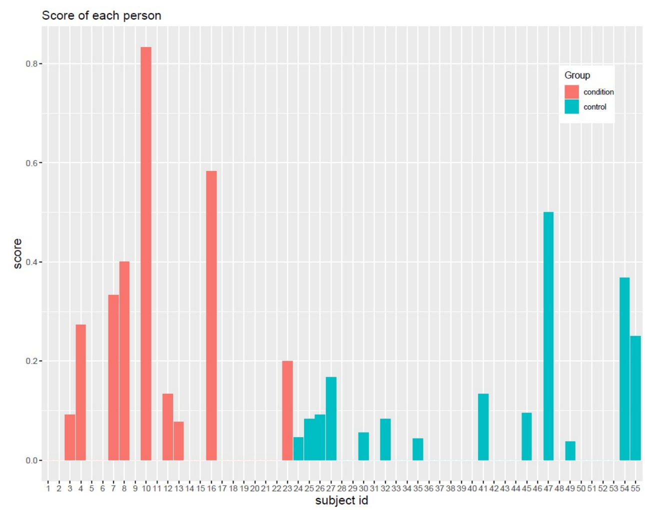
Correlation Graph



Clustering

Groups by score

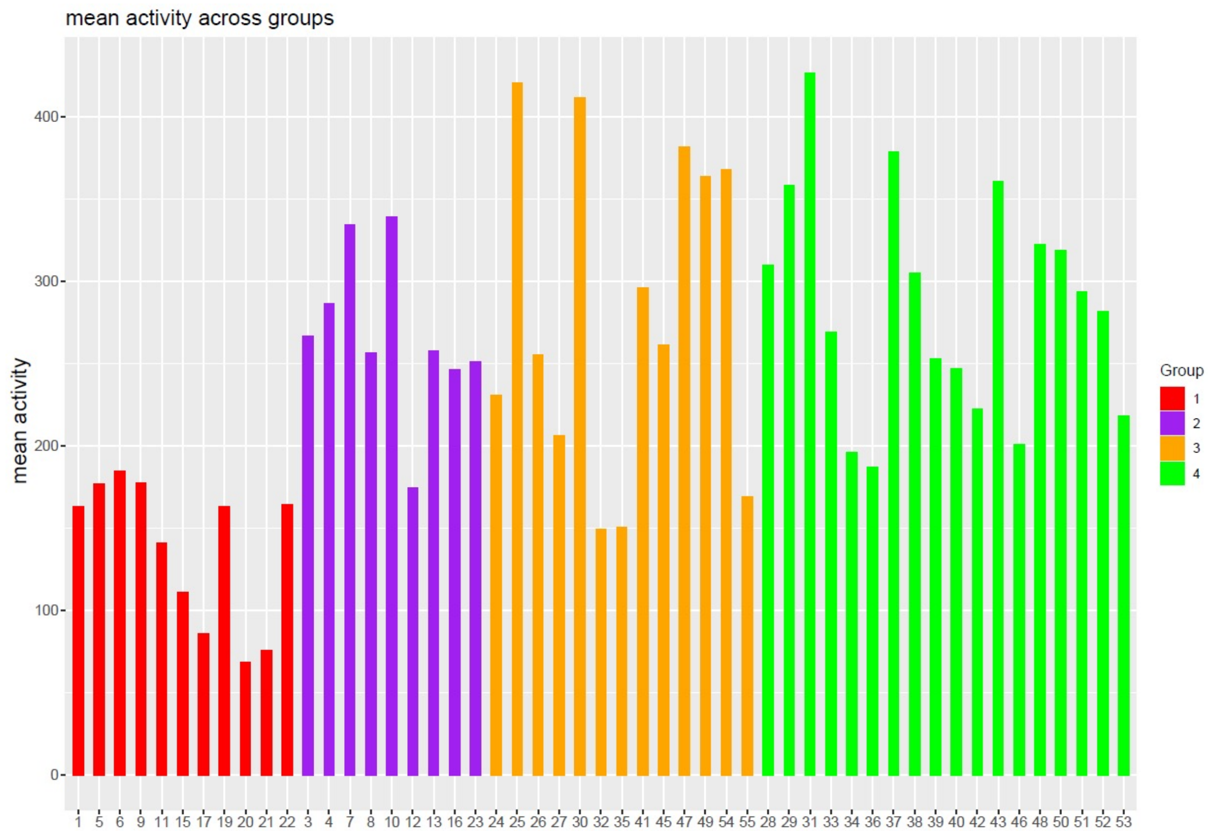




Group no	Original group	score	Color code
Group 1	Condition	0	Red
Group 2	Condition	>0	Purple
Group 3	Control	>0	Orange
Group 4	Control	0	Green

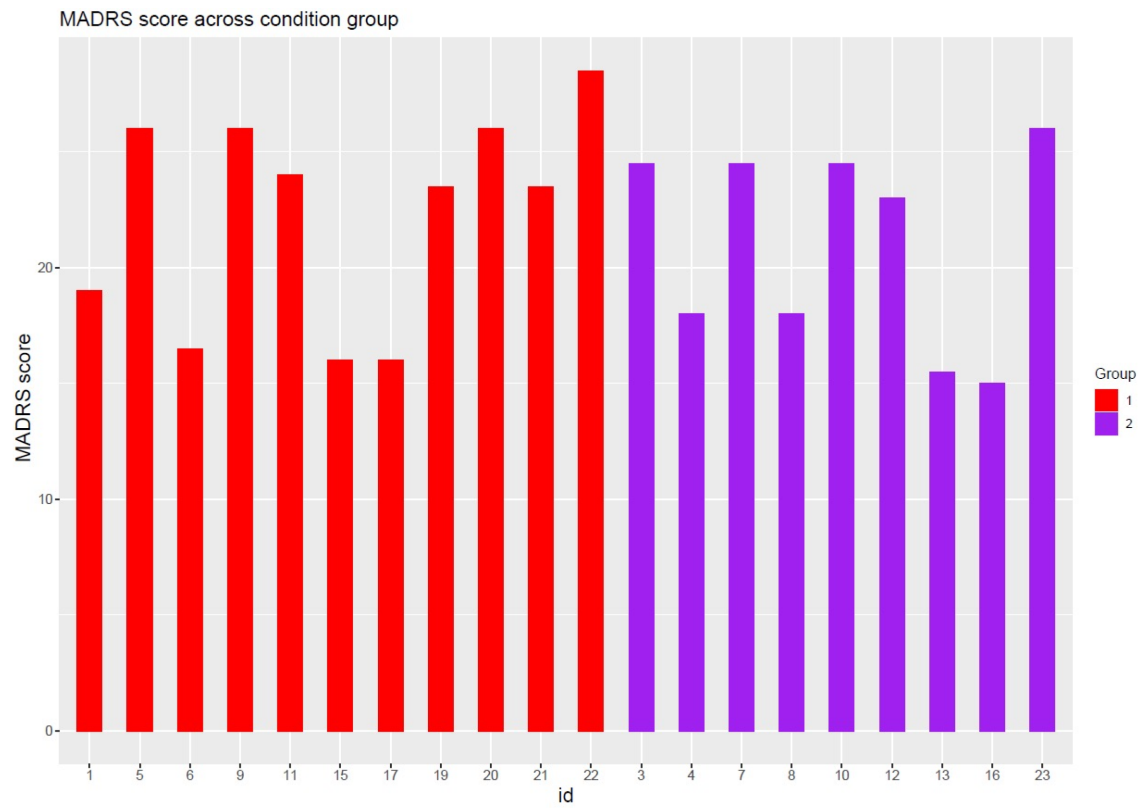
Mean activity across groups

- Red group have significantly lower mobility
- Mean activity represents their mobility and group



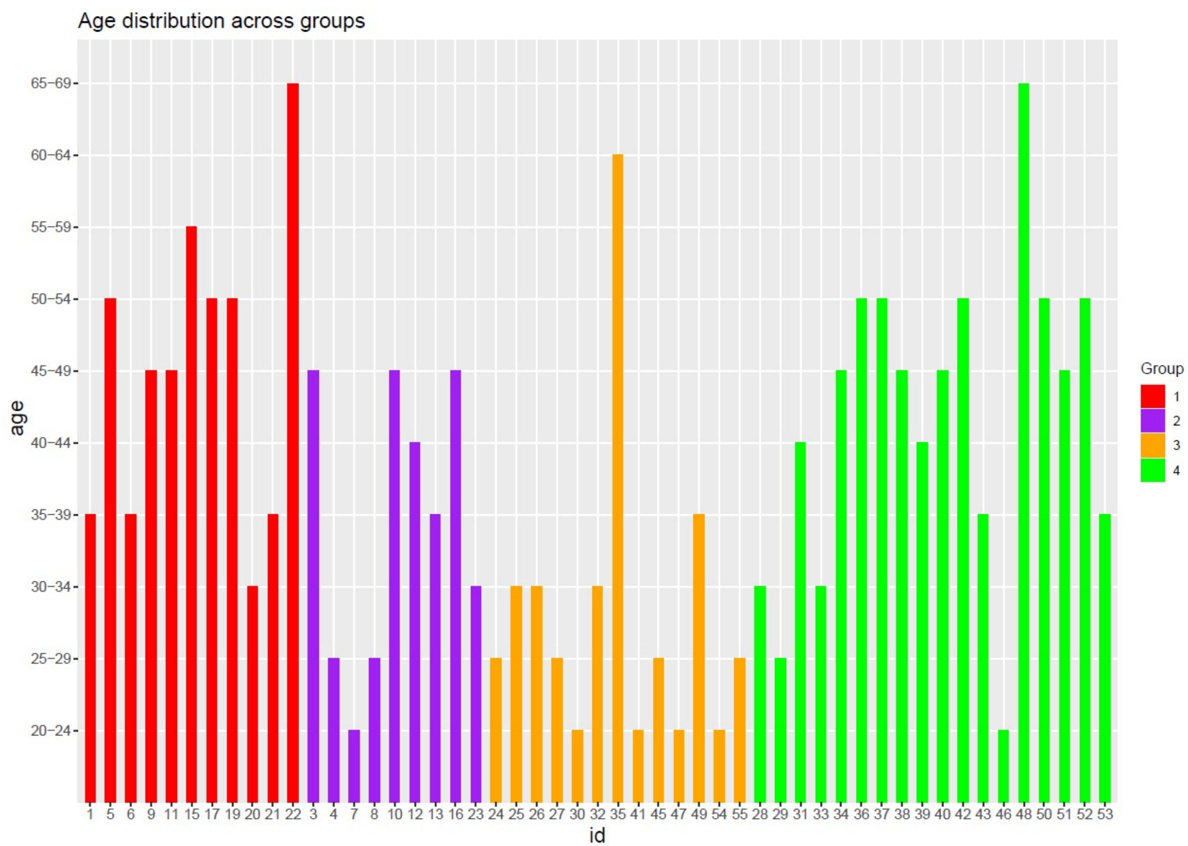
MADRS score for condition groups

- MADRS for red group is higher than purple group



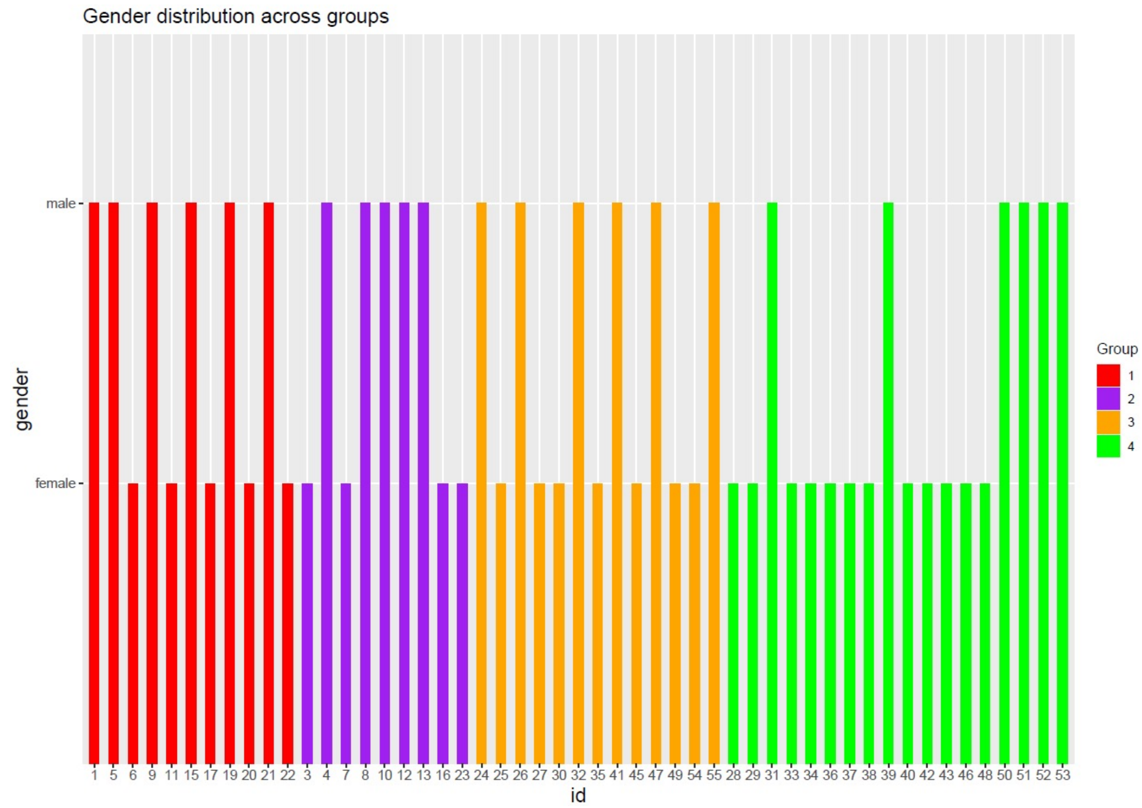
Age across groups

- All persons in red group > 30 years
- Majority of orange group are young participants < 30 years



Gender across groups

- Majority of green group are females



Discussion

- Person with id p7 belongs to condition & purple group
  - has high MADRS score but mean is also high
  - Age is 20-24. Youngest in condition group.
  - Regardless of depression his mobility skills are high, because he is Young?
- Singleton clusters – P2, P14, P18, and P44
  - Mobility of P14 & P18 is significantly lower
  - P2 – mobility is high but total number of days data collected is higher
    - 19 days for P2 while mean days of assessment is 12 days only
  - P44 has significantly lower number of days i.e 7 days
- Does these features cause them to clustered as singleton?

Conclusion

- Ability to identify groups with mobility as feature by utilizing population analysis-based Correlation network
- Identifying group of persons with similar mobility features without using class labels

Sciences at Crossroads

- Many Scientific disciplines are now at crossroads
- The proper penetration of IT represent tremendous challenges and great opportunities
- The importance of interdisciplinary approach and knowledge integration to problem solving
- The need for in-depth analysis and problem solving rather than the surface-level approaches
- Revolution is data collection requires a revolution in data Analytics – Complex Data demand Complex Methods

Acknowledgments

- **Bioinformatics Research Group**
  Kiran Bastola
  Sanjukta Bhoomwick
  Oliver Bonham-Carter
  Kate Dempsey
  Sachin Pawaskar
  Jay Peterson
  Ishwor Thapa
  Julia Warnke
  Sean West
- **Former Members of the Group**
  Alexander Churbanov
  Xutao Deng
  Huiming Geng
  Xiaolu Huang
  Daniel Quest
- **Biomedical Researchers**
  Steve Bonasera
  Steve Hinrichs
  Howard Fox
  Howard Gendelman
- **Wireless Sensors Research Group**
  Jon Youn
  Sangil Choi
  Abhilash Patlolla
  Elham Rastegari
  Ikhyun Youn
- **Funding Sources**
  NIH NIA
  NSF EPSCoR
  NSF STEP
  Nebraska Research Initiative