Data mining in Stabilometry: Application to patient balance study for sports talent mapping

Juan. A. Lara[1], David Lizcano[1], David[1] de la Peña, José M. Barreiro[2]

[1] **Open University of Madrid, UDIMA**

[2] **Technical University of Madrid**

udima
UNIVERSIDAD A DISTANCIA DE MADRID

Summary

- Introduction
- Background
- Data
- Methods
- Data analysis and results
- Discussion of Results
- Conclusions
- Future Work

Introduction (I)

- **Stabilometry** is a branch of medicine concerned with studying people's postural control
- Knowledge discovered from stabilometric data has led to quite a few advances in the field of medicine
- These data are not straightforward to analyse
- Specialized data mining techniques have to be used
- Classification models based on historical stabilometric data collected from different individuals

Introduction (II)

- Data include balance-related information (**independent variables** of our study) and other characteristics: *age, height and gender, as well as information related to the sports played by the respective individual* (each separately considered as a **dependent variable** in the conducted experiments)
- The resulting models explain these characteristics in terms of patient balance and …
- … act as a tool for recommending sports or sporting disciplines for young sports talents

Background

- We used a modern posturography device called *Balance Master* from NeuroCom® Internacional
- There are different types of **tests**
- Each test is divided into test **subtypes**
- As it is of special interest to the experts of this domain, we focused on the **US** (***Unilateral Stance***) test
- The aim of this test is to measure sway in patients standing on one foot with eyes open and with eyes closed

Data

- Time series data from a total of 56 individuals:
  - 15 professional basketball players
  - 18 elite ice skaters
  - 23 members of a control group of healthy people of different gender who are not professional sportspeople
- A framework was applied to acquire the raw data for analysis
- Indicators related to subject balance, such as:
  - sway velocity
  - number of recorded imbalances
  - number of recorded falls
  - sum of the lengths of the recorded falls
  - maximum intensity of the recorded fall measurements

Methods

- Data pre-processing tasks:
  - **T1**. *Age* attribute discretization and numeration
  - **T2**. *Height* attribute discretization and numeration
  - **T3**. *Sport* attribute numeration
  - **T4**. *Gender* attribute numeration
  - **T5**. *Sportsperson* attributization
  - **T6**. *Skater* attributization
  - **T7**. *Basketball Player* attributization
- Data mining task: **Classification**
- Data mining methods: **decision trees (CART)** and **logistic regression**

Data analysis and results (I)

- Summary of the experiments

			Instances for each class		
Experiment	**Subtype**	**Dependent variable**	**0 (%)**	**1 (%)**	**Total**
Exp1	Left	Age	62.50	37.50	56
Exp2	Right	Age	62.50	37.50	56
Exp3	Left	Gender	67.86	32.14	56
Exp4	Right	Gender	67.86	32.14	56
Exp5	Left	Height	60.71	39.29	56
Exp6	Right	Height	60.71	39.29	56
Exp7	Left	Sport	45.45	54.55	33
Exp8	Right	Sport	45.45	54.55	33
Exp9	Left	Sportsperson	58.93	41.07	56
Exp10	Right	Sportsperson	58.93	41.07	56
Exp11	Left	Skater	56.10	43.90	41
Exp12	Right	Skater	56.10	43.90	41
Exp13	Left	Basketball player	60.53	39.47	38
Exp14	Right	Basketball player	60.53	39.47	38

udima
UNIVERSIDAD A DISTANCIA DE MADRID

Data analysis and results (II)

- Results of the experiments in terms of classification accuracy (10-fold cross validation)

Experiment	Decision Trees	Logistic Regression
Exp1	55.36%	58.93%
Exp2	69.64%	69.64%
Exp3	69.64%	80.36%
Exp4	60.71%	64.29%
Exp5	73.21%	85.71%
Exp6	66.07%	75.00%
Exp7	60.61%	**90.91%**
Exp8	81.82%	84.85%
Exp9	60.71%	48.21%
Exp10	39.29%	55.36%
Exp11	56.10%	65.85%
Exp12	56.10%	63.41%
Exp13	63.16%	76.32%
Exp14	47.37%	71.05%
MEAN	**61.41%**	**70.71%**

Discussion of results

- Logistic regression yields better results in 12 out of the 14 conducted experiments
- *Left* subtype yields better results in five out of the seven cases (*Gender, Height, Sport, Skater* and *Basketball player* attributes)
- The variables that appear to be most related to balance are *Height* and *Sport*
- The models output in Experiments 7 and 8 (*Sport* variable) could be said to be especially applicable for mapping out the career of young sports talents

Conclusions

- Stabilometric data can provide an enormous amount of information
- We described the experiments conducted based on the stabilometric data of a series of individuals
- Knowledge about existing relationships between people's balance and other characteristics: e.g., close relationship between people's balance and the sport that sportspeople play
- Models can be applied to determine the best sport or sports discipline for each subject depending on their balance: state programmes for capturing young sports talents

udima
UNIVERSIDAD A DISTANCIA DE MADRID

Future Work

- Broaden the sample in order to conduct a richer and more comprehensive analysis
- Other stabilometric tests
- Other sports apart from basketball and skating: to consider sports that are more alike (for example, compare basketball players with handball or volleyball players)
- Other techniques in order to analyse the correlations between the analysed data: clustering, plots, SVMs
- To study the relationship between balance and gender and check whether it holds for other datasets, tests, sports, etc.
- analyse the causality between balance and sports discipline

Thanks for your attention

QUESTIONS?

Juan Alfonso Lara Torralbo

juanalfonso.lara@udima.es