

АНАЛІЗ УЗГОДЖЕНОСТІ СУДДІВСТВА З ВИКОРИСТАННЯМ PYTHON НА ПРИКЛАДІ ЗМАГАНЬ ЗІ СПОРТИВНИХ ТАНЦІВ

Р. Селіверстов

*Львівський національний університет імені Івана Франка,
вул. Університетська, 1, Львів, 79000, e-mail: roman.seliverstov@lnu.edu.ua*

З використанням бібліотек Python розроблено програмне забезпечення для аналізу протоколів змагань зі спортивних танців. На прикладі конкретного змагання визначено рівень узгодженості суддівських оцінок, його залежність від багатьох параметрів. Запропоновано можливі напрями удосконалення моделі.

Ключові слова: експертне оцінювання, коефіцієнт конкордації, Python.

1. ВСТУП

Інформаційні технології для виявлення закономірностей, формування зв'язків, отримання висновків на підставі аналізу різноманітних даних застосовують в усіх галузях, зокрема й у спорті. У багатьох видах спорту результат спортсмена (команди) можна визначити беззаперечно за кількістю набраних очок, часом, відстанню тощо. У таких випадках результат автоматично фіксується високоточними приладами або візуально (на підставі правил, які трактуються однозначно) суддями. Якщо виникають певні сумніви – завжди можна звернутися за допомогою до технічних засобів (відеоповтор, фотофініш, система фіксації взяття воріт тощо). Водночас у змаганнях з художньої гімнастики, стрибків у воду, фігурного катання, спортивних танців тощо, критерії оцінювання ґрунтуються на якісних показниках і замість (на додаток до) інструментальних методів використовують евристичні (інтуїтивні), основою яких є думка експертної комісії (суддівської бригади). Оскільки вимірними засобами під час оцінювання виступів спортсменів у змаганнях з таких видів спорту є люди, то через можливу неприязнь, корпоративні інтереси, фінансову зацікавленість та різноманітні емоційно-психологічні чинники постає проблема об'єктивності оцінювання.

Питання об'єктивності та кваліфікованості суддівства у спорті постійно піднімаються на чи після змагань вищого гатунку. Вони є найбільш значущими й актуальними для змагань любительського рівня, особливо для дитячих змагань, де помилка чи усвідомлене спотворення результату не має значного суспільного відголосу, але негативно впливає на морально-психологічний стан юних спортсменів. Тому діяльність суддів варто постійно контролювати, а для запровадження санкцій для несумлінних представників арбітражу потрібне наукове обґрунтування їхньої некомпетентності чи заангажованості. Стосовно спортивних танців, які є об'єктом цього дослідження, оскільки результати більшості змагань відкриті для широкого загалу, то існує можливість провести зовнішній аналіз узгодженості думок суддівських бригад разом з вірогідним внутрішнім аналізом (конкретної інформації про методи контролю вітчизняними федераціями й асоціаціями спортивного танцю за дотриманням суддями етичних норм у відкритих джерелах немає).

Для аналізу узгодженості суддівських оцінок у спортивних танцях зазвичай достатньо обчислювально-простих кількісних методів експертного оцінювання, які ґрунтуються на рангових коефіцієнтах конкордації Кендала та кореляції Спірмена. Застосування кваліметричних експертних методик у тренувальному процесі та спортивному арбітражі розглядали, наприклад, у працях [1, 2]. Щоправда, наведені у такого типу роботах результати обмежуються рекомендаціями та прикладами на основі гіпотетичних заздалегідь структурованих даних невеликого обсягу, отримання й аналіз яких не потребує залучення інформаційних технологій (принаймі достатньо офісного редактора електронних таблиць), або висновками щодо суддівства окремих змагань без опису моделі та методів.

Ми розробили й апробували програмне забезпечення для збирання, структурування та аналізу (з'ясування ступеня узгодженості суддівських оцінок) даних про результати фінальних стадій (дофінальні стадії оцінюють за іншою системою, що потребує відмінних від поданих у цій праці підходів до збирання й обробки інформації) змагань зі спортивних танців. Мовою програмування обрали Python через його відкритість і затребуваність у статистичній аналітиці та роботі з даними [3, 4, 5].

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Хоча в межах цього дослідження розглядаються невеликі обсяги даних, сформулюємо задачу у вигляді процесу data science, який складається з таких шести етапів [3].

1. Визначення мети дослідження. Не зменшуючи загальності, окреслимо мету так: зробити висновок про узгодженість оцінок членів суддівської бригади на окремо взятому турнірі зі спортивних бальних танців.

2. Збирання даних. Для подальшого аналізу потрібні дані про вікову категорію та клас учасників фіналу, програму й оцінки (ранги), поставлені суддями за кожен танець програми.

3. Підготовка даних. Подання даних у csv-файлі, структура якого виглядає так:

Category, Class, Dance, Marks

$(ct)_i, (cl)_i, (dn)_i, (mr)_i$.

Тут вжито такі позначення: $i = 1, 2, \dots, n$ – порядкові (ідентифікаційні) номери фіналів; n – загальна кількість фіналів; ct_i – вікова категорія (діти, ювенали, юніори, молодь, ...); cl_i – клас спортсмена (школа, початківець, E, D, ...); dn_i – танець (повільний вальс, танго, ча-ча-ча, ...);

$$mr_i = x_{11}^i x_{12}^i \dots x_{1n_i}^i - x_{21}^i x_{22}^i \dots x_{2n_i}^i - \dots - x_{m_i 1}^i x_{m_i 2}^i \dots x_{m_i n_i}^i$$

– рядок (string) суддівських оцінок, де $x_{jk}^i \in 1, 2, \dots, m_i$ – ранг (місце), присуджений у i -му фіналі j -му учаснику (танцювальній парі) k -м суддею; m_i – кількість учасників фіналу; n_i – кількість членів суддівської бригади.

4. Дослідження даних. Перевірка коректності даних, яка полягає в усуненні фізично неможливих і неіснуючих значень, хибодруків, помилок інтерпретації, надлишкових пробілів, розбіжностей у регістрі символів та ін. Крім того, має

виконуватись умова, яка унеможливило б присудження суддею того самого місця кільком учасникам:

$$\forall i, j, k, l \in \mathbb{Z} (1 \leq i \leq n, 1 \leq j \leq m_i, 1 \leq k, l \leq n_i) : l \neq k \Rightarrow x_{jk}^{(i)} \neq x_{jl}^{(i)}.$$

5. Моделювання даних. Обчислення коефіцієнтів конкордації та перевірка його значущості за критерієм узгодженості Пірсона. Оскільки коефіцієнт конкордації належить до класичних мір статистичної залежності, то формул для його обчислення та перевірки значущості не наводимо.

6. Відображення й автоматизація. Висновок про рівень узгодженості думок суддів та існування узагальнених групових поглядів серед членів суддівських бригад залежно від параметрів задачі (вікової категорії, класу, програми тощо). Загальний висновок про якість суддівства та рекомендації щодо його поліпшення.

3. ЛІСТИНГИ ПРОГРАМ І РЕЗУЛЬТАТИ ВИКОНАННЯ

Прикладом для аналізу суддівських оцінок слугуватимуть відкриті класифікаційні змагання зі спортивних танців “The legend of Golden Lviv”, які проводила АСТУ (Асоціація Спортивного Танцю України) у м. Львів, 4 листопада 2017 року, результати яких є у вільному доступі на танцювальному ресурсі flymark.com.ua. Щоб полегшити подальшу обробку даних (містяться в архіві lgl_protocols.zip з doc-файлами) доцільно перетворити їх у txt-формат (див. лістинг 1). Для цього використана програма для конвертування файлів catdoc, яку у нашому випадку скрипт Python запускає з командної оболонки Linux.

Лістинг 1. Перетворення даних у текстовий формат

```
from zipfile import ZipFile
from os import system
with ZipFile('lgl_protocols.zip') as myzip:
    files = myzip.namelist()
    myzip.extractall()
for item in files:
    system('catdoc {0} > {1}.txt'.format(item, item[:-4]))
```

Лістинг 2 описує процедуру формування csv-файлу, а лістинги 3-5 – функції, які виймають з текстових файлів потрібні для подальшого аналізу дані про загальні характеристики фіналу (вікову категорію та клас танцювальних пар), а також оцінки (ранг пари у фіналі) від кожного судді за кожен танець. Лістинги, які стосуються перевірки коректності даних, тут не наводяться.

Лістинг 2. Формування csv-файлу з даними

```
from os import listdir
txtfiles = filter(lambda x: x.endswith('.txt'), listdir())
data = open('data.csv', 'w')
s = 'Category,Class,Dance,Marks\n'
for item in txtfiles:
    txtfile = open(item, 'r')
    lines = txtfile.readlines()
    ct = get_category(lines[11].strip())
    cl = get_class(lines[11].strip())
    final = get_marks(lines)
    for dn in final.keys():
        mr = final[dn][-1]
        s += '{0},{1},{2},{3}\n'.format(ct, cl, dn, mr)
    txtfile.close()
data.write(s[:-1])
data.close()
```

Лістинг 3. Пошук вікової категорії

```
def get_category(text):
    index_1 = text.find(' ')
    if text[0] == 'Ю':
        index_2 = index_1 + text[index_1+1:].find(' ')
        text = text[:index_2+1]
        if '1' in text or '2' in text:
            return text[:index_2+1]
        else:
            return text[:index_1] + ' 1+2'
    return text[:index_1]
```

Лістинг 4. Пошук класу

```
def get_class(text):
    classes = ['B', 'C', 'D', 'E', 'H', 'Rising Stars']
    endindex = text.find(',')
    text = text[:endindex]
    for item in classes:
        if ' '+item in text:
            return item
```

Лістинг 5. Пошук фінальних оцінок

```
def get_marks(lines):
    marks = {}
    for line in lines:
        if '(' in line and '\n' in line:
            dance = line[:line.find('(')-1]
            marks[dance] = ''
            continue
        if len(line)>10:
            s = (line[:-1]).replace(' ', '')
            if s.isdigit():
                marks[dance] += '{0}-'.format(s)
    return marks
```

Щоб ліпше зрозуміти наведені лістинги, наводимо типові і водночас унікальні за структурою рядки txt-протоколів з необхідними даними (вікова категорія і клас завжди в 11-му рядку):

```
...
Юніори 2 E, LA
...
Ча-ча-ча (C)
...
6 6 3 2 5 4 5 1 4 7 2 5 1 3 1
...
7 4 5 6 4 7 8 4 1 5 5 3 8 4 3
...
Самба ( S )
...
```

Результатом виконання коду з лістинга 2 є файл data.csv зі 137 записами (кількість оцінених танців у всіх фіналах разом). Зазначимо, враховували лише виступи танцювальних пар, сольні виступи не враховували. Лістинг 6 демонструє обчислення коефіцієнтів конкордації (враховуючи перевірку їхньої значущості) та деяких допоміжних даних (розмір суддівської бригади, кількість учасників фіналу, програма (стандартна або латиноамериканська)), а у лістингу 7 наведено зразки

(фрагменти коду) обчислення середніх коефіцієнтів конкордації згрупованих за певними ознаками записів.

Лістинг 6. Моделювання даних

```
import pandas
from numpy import mean
from scipy.stats import chi2
def get_marks(data):
    marks = data['Marks'].strip().split('-')
    return [list(map(int,item)) for item in marks]
def get_referies(data):
    return len(data['Ranges'][0])
def get_pairs(data):
    return len(data['Ranges'])
def get_style(data):
    LA = ['Ча-ча-ча', 'Джайв', 'Румба', 'Самба', 'Пасодобль']
    if data['Dance'] in LA:
        return 'LA'
    return 'ST'
def get_concordance(data):
    referies = get_referies(data)
    pairs = get_pairs(data)
    s_ranges = [sum(item) for item in data['Ranges']]
    mean_range = mean(sum_of_ranges)
    ssd = sum([(item - mean_range)**2 for item in s_ranges])
    return 12 * ssd / referies**2 / (pairs**3 - pairs)
def get_significance(data):
    x2p = data['W']*data['Referies']*(data['Pairs']-1)
    return round(1 - chi2.cdf(x2p,data['Pairs']-1),3)
data = pandas.read_csv('data.csv')
data['Style'] = data.apply(get_style, axis = 1)
data['Ranges'] = data.apply(get_marks, axis = 1)
data['Referies'] = data.apply(get_referies, axis = 1)
data['Pairs'] = data.apply(get_pairs, axis = 1)
data['W'] = data.apply(get_concordance, axis = 1)
data['P-value'] = data.apply(get_significance, axis = 1)
```

Лістинг 7. Приклади обчислення групових коефіцієнтів конкордації

```
print(data['W'].mean())
print(data.groupby(['Class'])[['Class', 'W']].agg([mean]))
col_to_show = ['Class', 'Category']
print(data.pivot_table(['W'], col_to_show, aggfunc='mean'))
```

Отримано 137 коефіцієнтів конкордації, з яких для рівнів значущості 0,05 значущими виявилися 110. Середнє значення коефіцієнта конкордації змагань – 0,36 (середня кількість суддів 11 і середня кількість учасників фіналу 7). Нижче подані значення коефіцієнтів конкордації для різних категорій, класів, танців і програм у форматі “Ознака: коефіцієнт конкордації (кількість записів)” та у вигляді зведеної таблиці (див. табл. 1):

Таблиця 1

Коефіцієнти конкордації

Категорія “Діти”:	0,23	(4)
Категорія “Ювенали 1”:	0,42	(10)
Категорія “Ювенали 1+2”:	0,44	(16)
Категорія “Ювенали 2”:	0,18	(10)

Категорія “Юніори 1”:	0,30	(28)
Категорія “Юніори 2”:	0,36	(28)
Категорія “Молодь”:	0,63	(10)
Категорія “Молодь+Дорослі”:	0,31	(21)
Категорія “Дорослі”:	0,47	(10)
Клас “Н”:	0,23	(20)
Клас “Е”:	0,30	(27)
Клас “D”:	0,22	(32)
Клас “Rising Stars”:	0,52	(48)
Клас “В”:	0,52	(10)
Танець “Повільний вальс”:	0,40	(19)
Танець “Танго”:	0,40	(14)
Танець “Віденський вальс”:	0,44	(10)
Танець “Квікстеп”:	0,39	(19)
Танець “Фокстрот”:	0,61	(5)
Танець “Ча-ча-ча”:	0,27	(20)
Танець “Самба”:	0,29	(15)
Танець “Румба”:	0,41	(10)
Танець “Джайв”:	0,30	(20)
Танець “Пасодобль”:	0,47	(5)
Стандартна (європейська) програма:	0,41	(67)
Латиноамериканська програма:	0,32	(70)

Таблиця 2

Зведена таблиця середніх коефіцієнтів конкордації для різних вікових категорій і класів спортсменів

	“Н”	“Е”	“D”	“Rising Stars”	“В”
“Діти”	0,23	–	–	–	–
“Ювенали 1”	0,23	0,55	–	–	–
“Ювенали 1+2”	–	–	0,32	0,55	–
“Ювенали 2”	0,18	0,17	–	–	–
“Юніори 1”	0,21	0,34	0,23	0,37	–
“Юніори 2”	0,30	0,24	0,20	0,59	–
“Молодь”	–	–	–	0,63	–
“Молодь+Дорослі”	–	0,09	0,13	–	0,52
“Дорослі”	–	–	–	0,47	–

4. ВИСНОВКИ

Загалом напрошується висновок про недосконалість суддівства. Про це свідчить низьке значення середнього по турніру коефіцієнта конкордації і великий розкид значень – від 0,01 до 0,83. Лише 4 виступи (3%) були оцінені з високим рівнем узгодженості (коефіцієнт конкордації вищий за 0,7), тоді як узгодженість 64 виступів (47%) незадовільна (коефіцієнт конкордації нижчий за 0,3).

Проте проблемнішою вбачається тенденція до зростання узгодженості оцінок із зростанням класу (трохи менш яскраво виражено – віку) спортсменів (і навпаки, чим молодші спортсмени і нижчий клас змагання, тим гірша ситуація з суддівством), яку можна спостерегти, крім поданих вище відповідних групових середніх коефіцієнтів конкордації, за такими фактами:

а) з 30 найбільш узгоджено оцінених виступів 29 стосуються високих класів “В” і “Rising Stars”, 25 – старших вікових категорій “Юніори 2”, “Молодь” та “Дорослі”;

б) усі 30 найменш узгоджено оцінених виступів стосуються низьких класів “Н”, “Е” та “D”;

в) візуально видно, що коефіцієнти конкордації у правому нижньому куті (див. табл. 2) зазвичай більші за значення з лівого верхнього кута.

Вплив конкретного танцю майже не впливає на узгодженість суддівства, хоча чомусь середній рівень узгодженості оцінювання танців латиноамериканської програми нижчий порівняно зі стандартною (європейською) програмою.

Звісно, проведене дослідження не ґрунтовне. Детальніший аналіз можна провести, наприклад, на підставі пошуку узагальнених групових поглядів, з’ясування міри впливу кожного судді на узагальнену узгодженість оцінок, порівняння показників узгодженості суддівських оцінок по кожному учаснику фіналу, визначення інформаційної міри збігу думок тощо [6, 7]. Можна заглиблюватися далі, враховуючи психоемоційні й етичні чинники. Наприклад, у моделі немає часової компоненти, тобто не досліджено ефект втомлюваності суддів, адже нерідко змагання відбуваються впродовж двох днів, починаючись зранку і закінчуючись пізно вночі. Не ставилося за мету робити аналіз “я тобі – ти мені”, адже майже кожен суддя одночасно є тренером танцювальної пари (пар), які беруть участь у тих самих змаганнях. На підставі виконаного аналізу можна зробити висновки стосовно об’єктивності результатів, які можуть слугувати підґрунтям для прийняття рішень, які сприятимуть якісному підвищенню рівня суддівства.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Костюкевич В.М.* Теоретико-методичні основи контролю у фізичному вихованні та спорті: монографія / В.М. Костюкевич, Є.П. Врублевський, Т.В. Вознюк та ін.; за заг. ред. В.М. Костюкевича. – Вінниця: ТОВ “Планер”, 2017. – 191 с.
2. *Бакулина Е.Д.* Совершенствование объективности арбитражной системы и ее влияния на критерии оценок в художественной гимнастике / Е.Д. Бакулина // Ученые записки РГСУ (Приложение), 2006. – № 2. – С. 56-63.
3. *Cielen D.* Introducing Data Science. Big data, machine learning, and more, using Python tools / D. Cielen, A. Meysman, M. Ali. – New-York: Manning Publications, 2016. – 320 p.
4. *Sweigart A.* Automate the boring stuff with Python. Practical programming for Total Beginners / A. Sweigart. – San Francisco: No Starch Press, 2015. – 480 p.
5. *Ketkar N.* Deep learning with Python. A hands-on introduction / N. Ketkar. – New-York: Apress, 2017. – 169 p.
6. *Грабовецький Б.Є.* Планування та економічне прогнозування: навчальний посібник / Б.Є. Грабовецький. – Вінниця : ВНТУ, 2013. – 66 с.

7. Новосад В.П. Методологія експертного оцінювання: конспект лекцій для використання в навчальному процесі в системі підвищення кваліфікації кадрів / В.П. Новосад, Р.Г. Селіверстов. – Київ: НАДУ, 2008. – 48 с.

Стаття: надійшла до редколегії 26.09.2018

доопрацьовано 10.10.2018

прийнята до друку 31.10.2018

PYTHON-ANALYSIS OF REFEREEING CONCORDANCE BY THE EXAMPLE OF BALLROOM DANCE COMPETITIONS

R. Seliverstov

Ivan Franko National University of Lviv,

Universytetska Str., 1, Lviv, 79000, e-mail: roman.seliverstov@lnu.edu.ua

With the use of Python libraries the software for analysis of ballroom dance score-sheets is developed. On the example of particular competition the level of refereeing concordance and its dependence on a number of parameters are determined. Some possible directions for improvement of the model are proposed.

Key words: expert estimation, Kendall's tau coefficient, Python.