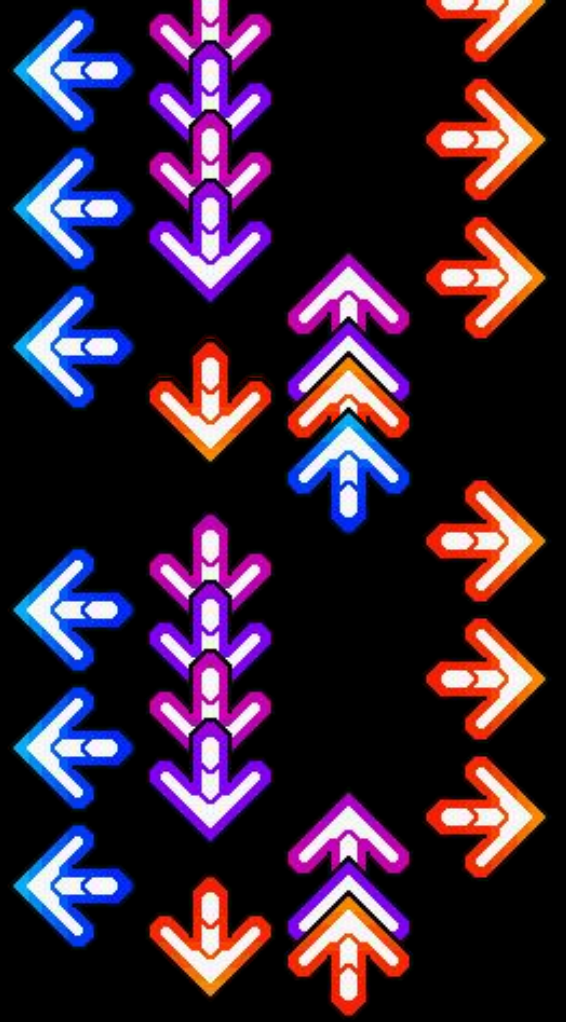
But We Regress

Predicting StepMania Song Difficulty Using Regression

Svitlana Glibova

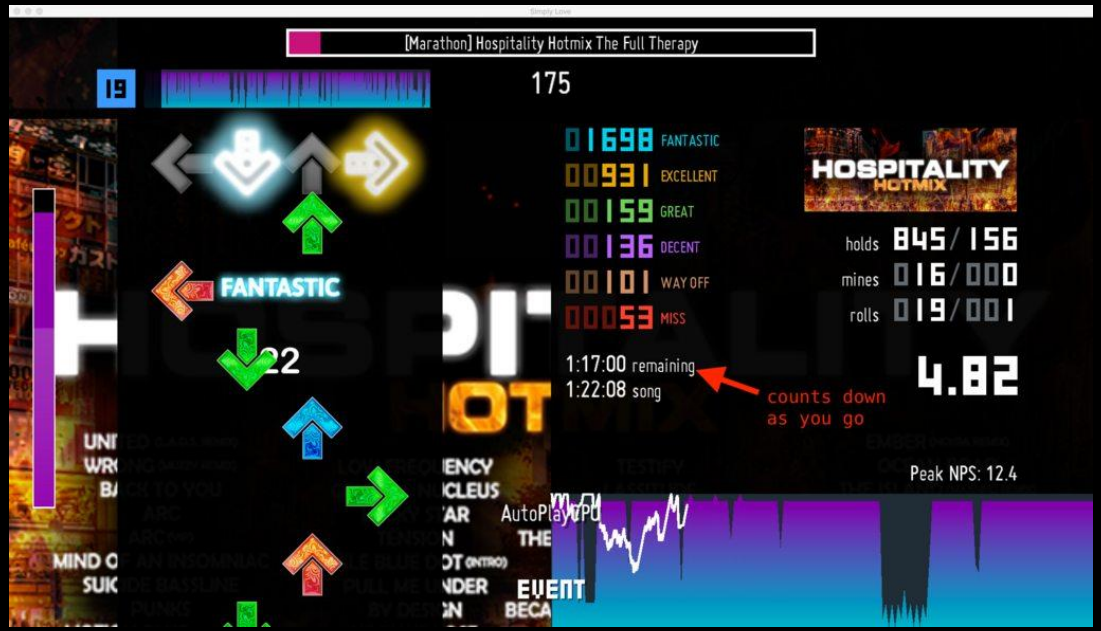
outline

- What is StepMania?
- Data Collection
- Modeling
- Results
- Next Steps



What is StepMania?

- Open-source software for rhythm games
- Endlessly customizable



Credit: [SimplyLove by quietly-turning and hurtpiggypig]

Data Collection

ITGPACKS.COM - community-driven spreadsheet

Chart Parser built by Tim Murphy, fellow dance-gamer and software engineer

3200 Non-Stamina
2250 Stamina

Example features collected:

- Step Count
- BPM weighted average
- NPS - notes per second
- Technical elements

Technical vs. Stamina Play

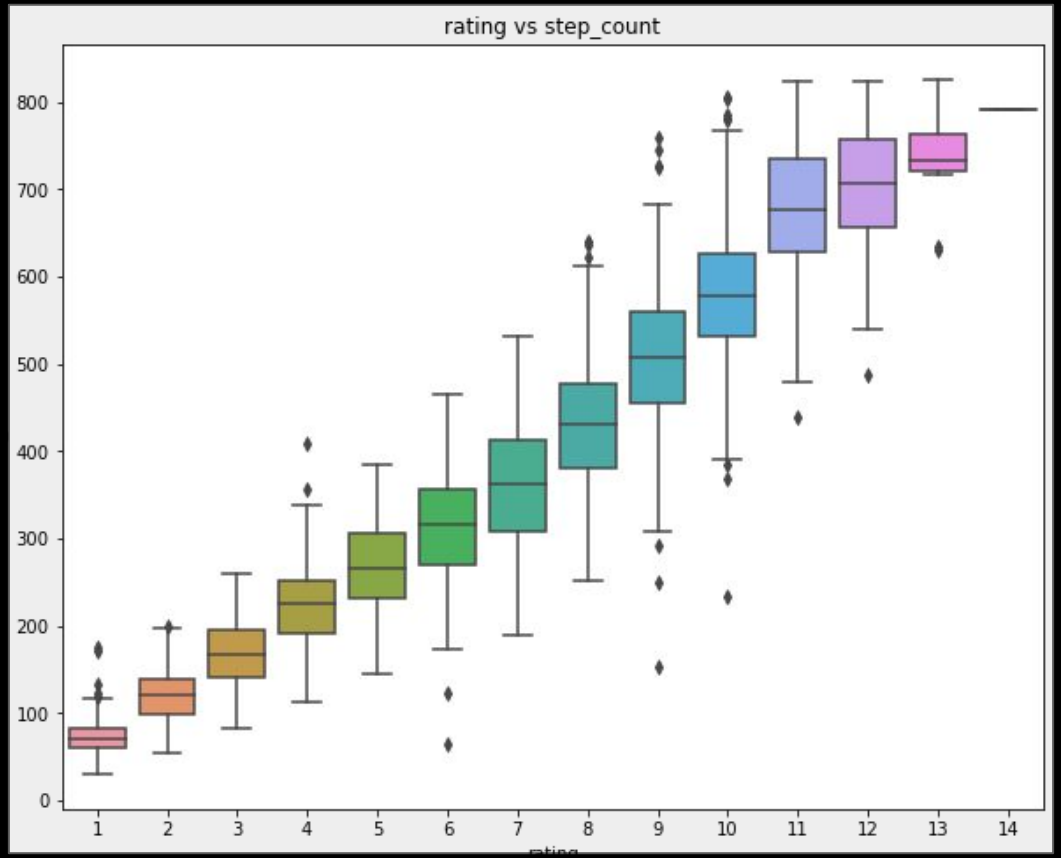
Technical

- More pattern complexity
- Shorter songs
- More body movement overall

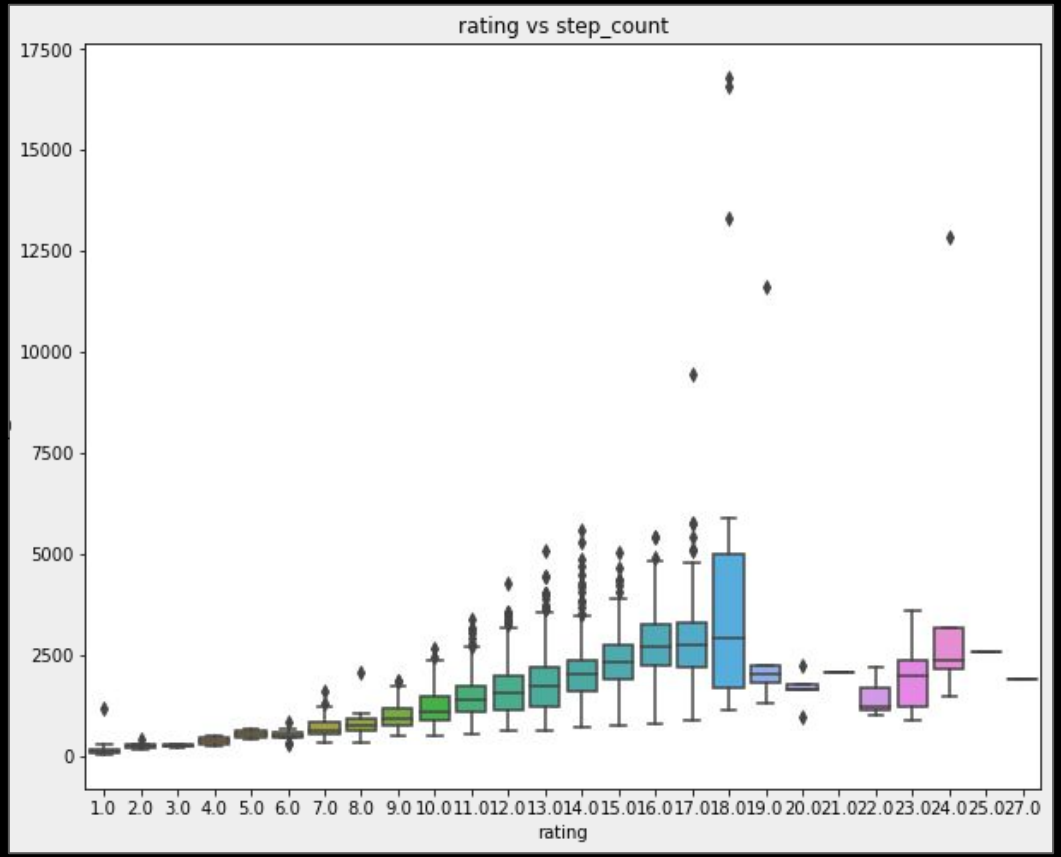
Stamina

- Less technical patterns
- Longer songs
- Tests endurance/consolidation

Not-Stamina


rating vs step_count
800
700
600
500
400
300
200
100
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Stamina

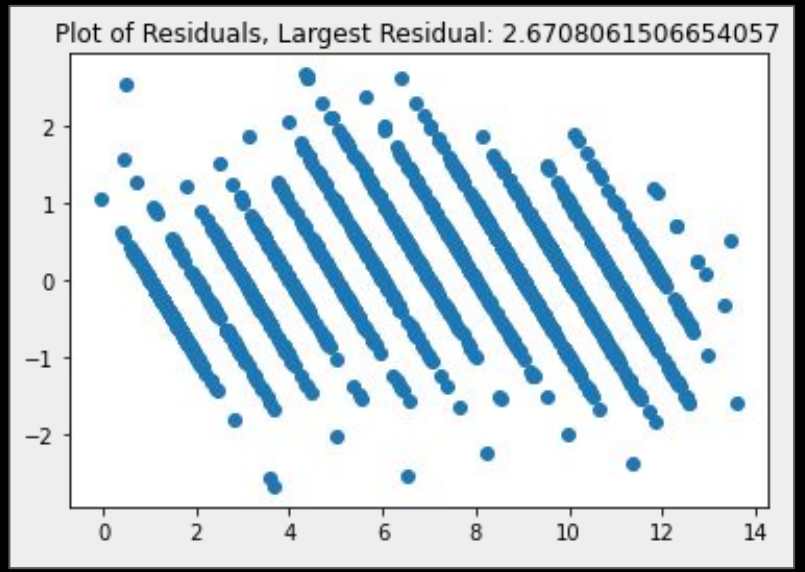

rating vs step_count
17500
15000
12500
10000
7500
5000
2500
0
1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0 13.0 14.0 15.0 16.0 17.0 18.0 19.0 20.0 21.0 22.0 23.0 24.0 25.0 27.0
rating

First Simple Model - Technical

Linear Regression

R-squared: .951

MSE: .642

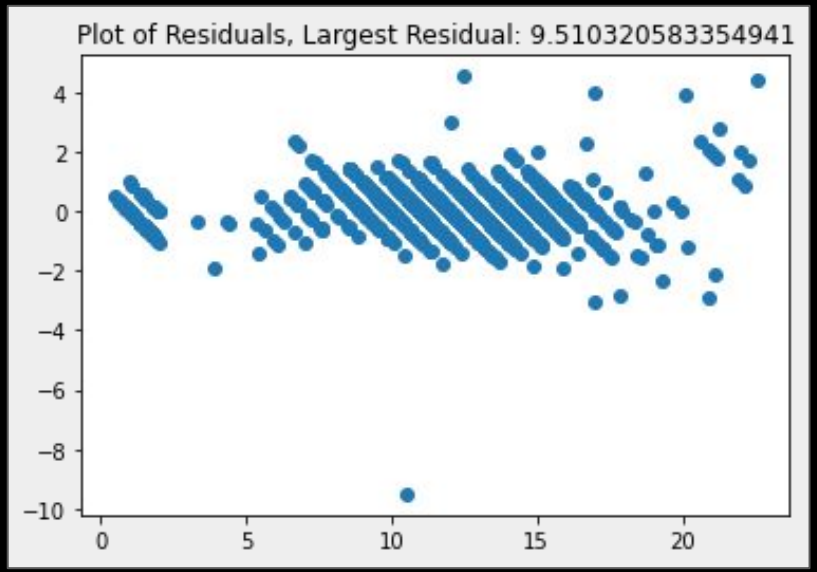


First Simple Model - Stamina

Linear Regression

R-squared = .958

MSE = .502

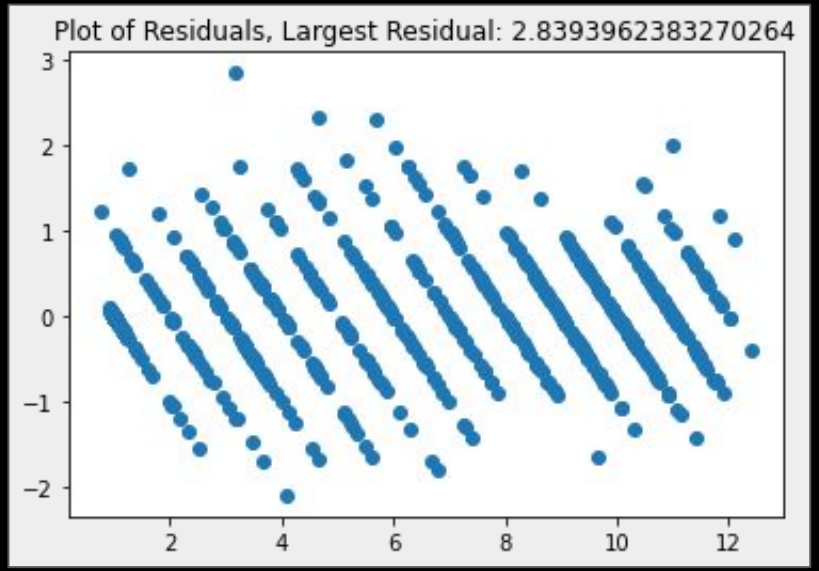


Final Model - Technical

XGBoost Regressor

R-squared = .968
MSE = .404

important features:
.575 NPS/measure avg.
.29 Song NPS

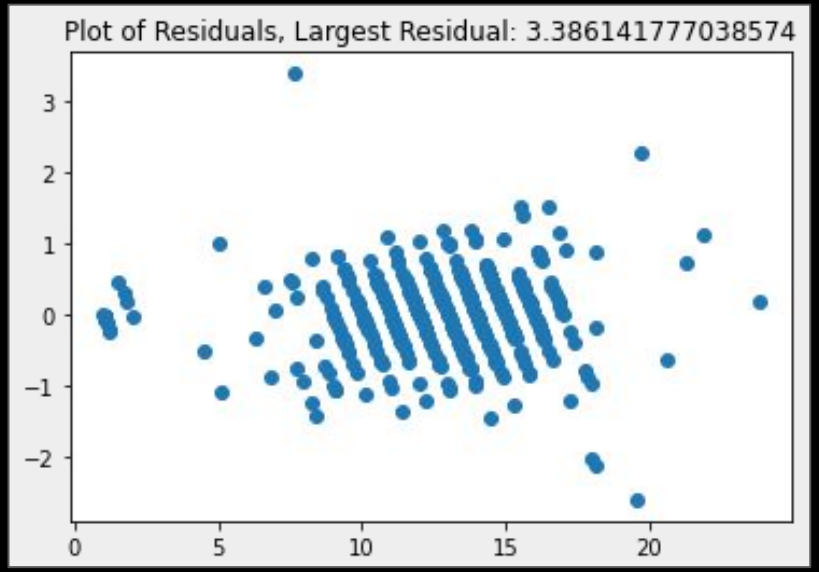


Final Model - Stamina

XGBoost Regressor

R-squared = .979
MSE = .282

important features:
.546 NPS/measure avg.
.229 NPS/measure max.



Next Steps

- Deployment as a Flask application
- Tuning a classification model
- Further data acquisition

THANK YOU!

svitlana glibova
s.glibova@gmail.com |  |