

Using learning analytics to analyze learning outcomes in a serious game

Ángel Serrano Laguna



UNIVERSIDAD
COMPLUTENSE
MADRID



UCM
e-learning group

Serious games

~~Educational videogames~~

Serious games

~~Educational videogames~~

~~Serious games~~

Applied games

Videogames as educational tools

- The **GOOD:**

- They can represent real settings (that can be dangerous or expensive in real life)
- Enable learning by doing
- Usually more engaging than other activities

- The **BAD:**

- They can be expensive to develop
- Transforming learning process into game mechanics can be difficult
- Low adoption



Can we use serious games as reliable
assessment tools?

The problem

- The vast majority of serious games use **written test/exams to validate learning** in serious games
 - (Pre-Test) - Serious game - Exam about content in serious game
- Our hypothesis:
 - Predict exam results analyzing interactions with serious games
- The goal
 - Get rid of the exam and convert the serious game in an **automatic assessment tool**

Case study: A serious game analysis

- A serious game producing interactions
- An exam assessing players learning
- The goal: predict exam results from serious game interactions
- The way: Machine Learning

About the game

- The game represent the story of a classical theatre play
- 3rd person adventure game (point & click)
- PC Game
- Educational Goals:
 - Students should learn the main plot points of the play
 - Students should learn several language concepts
- Duration: 30 minutes
- Game design
 - Plot knowledge was implicit during the gameplay
 - Language knowledge was associated to small mini-games
 - Each language concept appeared at least twice: once as practice and once as testing

The experiment

- 305 high school students
- 12-16 years old
- 8 high schools in Madrid area
- Experiment
 - (Written Pre-test) (5 minutes)
 - Serious game play (30 minutes)
 - Written Exam (5 minutes)

Data collected

- Serious game
 - Variables
 - Meaningful and not meaningful
 - Times
 - In each scene
 - In conversations with characters
 - Choices
 - In conversations
 - In game exploration
 - Interactions
 - Clicks

339 data points per student

- Exam (8 questions)
 - 4 points plot knowledge
 - 4 points language knowledge

8 data point per student :

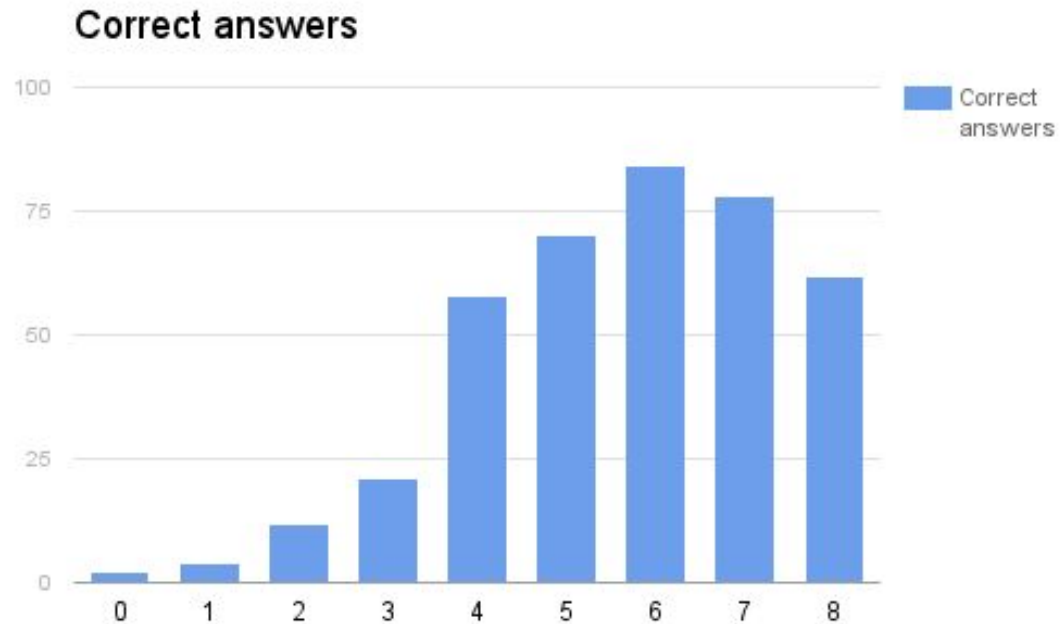
1 data point with final mark

First attempt

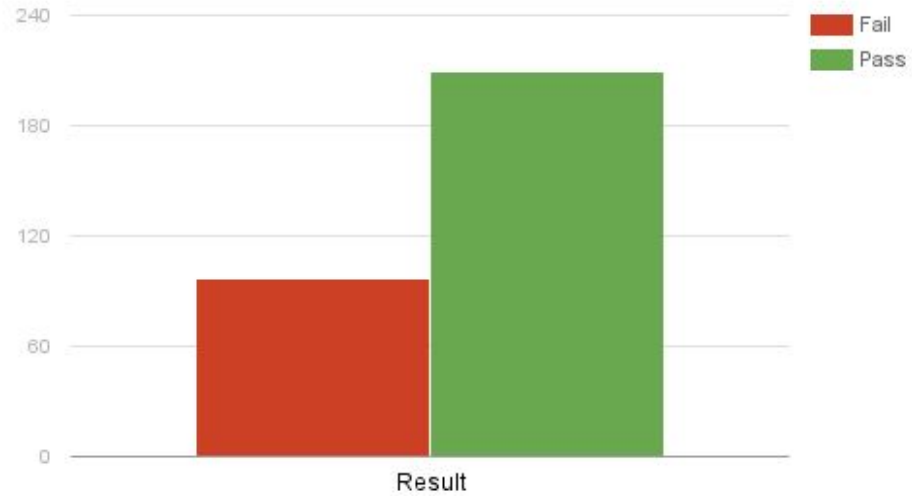
- 339 features per player => 1 output (a value between 0-8)
- Accuracy:
 - Players whose prediction is in the ± 1 range of the actual score
 - Cross-validation
- Machine learning techniques
 - Regression
 - Classification
- Algorithms
 - Naïve Bayes
 - Support Vector Machines
 - Decision Trees
 - ...

21%

Unbalanced data



Simplifying the problem



Choosing significant features

- Those features that have the biggest differences among groups, for example:
 - A boolean game variable is significant if FALSE is the most frequent value in one group and TRUE is the most frequent in the other
 - High difference in mean values in continuous variables (times, scores...)
 - High variability in options selected among groups

New attempt

- 23 features per player -> 1 output (the class 0: FAIL or 1 PASS)
- Accuracy:
 - % of players correctly categorized in FAIL and PASS
- Machine learning techniques
 - Classification
- Algorithms
 - Naïve Bayes
 - Support Vector Machines
 - Decision Trees
 - ...

86% - 66%

PASS - FAIL

Can we get rid of the exam and use only the serious game to assess students?

Conclusions

- The long term goal seems legitimate: we could use serious games as assessment tools
- There are a lot possibilities trying to compare assessment from traditional and non-traditional assessment methods
- Machine learning is not perfect
 - But there is not a perfect solution (probably)

Thanks!
