## An Expert System that Involves Students in Evaluation Process

Ibrahim A. Hameed, PhD, Associate Professor<br>Nestleder for forskning og innovasjon<br>Institte of IKT og realfag (IIR)<br>ibib@ntnu.no<br>https://www.ntnu.edu/employees/ibib

## LÆRINGSFESTIVALEN

## Every exam story

- Exam consists of 5 questions.
- Teacher assigned 20 marks for each question

- Student A solved questions 1, 2, 3, and 4
- Student B solved questions 2, 3, 4, and 5
- Total grade of $A=$ total grade of $B$
- It happens that question 1 is more complex and more important to the curriculum
- The question now:
$-A=B$ ? (traditional evaluation system)
$-A>B$ ? (new evaluation system)


## Objective

- Provide fairness to evaluation.
- Build en evaluation system that can consider other factors such as complexity, importance and difficulty of exam questions.
- Involve students in the evaluation process by allowing system to vote for complexity and importance.
- Use students' votes to alter/adjust marks assigned the teacher (if necessary).
- Put the system online and make it available for teachers and students in Norway and around the world.
- Evaluate if such as a system could positively impact education and learning in our schools.


## System block diagram


voting

## System block diagram (2)



## The inputs to the system

- Accuracy rate matrix: is obtained after correcting and grading answer sheets (by teacher)
- Answer time matrix: is easily obtained for computerbased exams
- Importance vector: a number between 0 and 1 for each question (by various domain experts such as teachers and students)
- Complexity vector: a number between 0 and 1 for each question (by various domain experts such as teachers and students)


## Fuzzy Inference system (FLS)



| Accuracy |  |  |
| :---: | :---: | :---: |
|  | Very low |  |
| Very low | Medium |  |
| Low | Low | . |
| Medium | Low |  |
| High | Very low |  |
| Very high | Very low |  |

Time rate
1
1

| Difficulty/Answer- <br> cost | Complexity/mportance |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | Very low | Low | Medium | High | Very high |
| Very low | Very low | Very low | Low | Low | Medium |
| Low | Very low | Low | Low | Medium | High |
| Medium | Low | Low | Medium | High | High |
| High | Low | Medium | High | High | Very high |
| Very high | Medium | High | High | Very high | Very high |

## Strawberry and apple

- Size (radius)
- Small \& large
- Rule 1: If radius is small THEN fruit is strawberry
- Rule 2: If radius is large THEN fruit is apple


## membership



Fuzzify


## Involvement of students in evaluation process

(1) Voting for complexity and importance ratios: effective ratios are the average of votes
(2) Voting for membership function shapes and distribution

## Membership functions' shape

Student 1

Student $n$

Effective
(average)


## EduEval: online tool

## Test Exam 2

```
Test Exam 2 for ID
101010 (Change?)
Number of questions:
5
```

Number of
evaluations: 2

## Current weights:

| Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Sum |
| :--- | :--- | :--- | :--- | :--- | :--- |
| 20 | 20 | 20 | 20 | 20 | 100 |

Re-evaluated weights

| Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Sum |
| :--- | :--- | :--- | :--- | :--- | :--- |
| 19.79 | 19.27 | 21.89 | 19.79 | 19.27 | 100 |

Upload Accuracy Matrix

## 1) Teach dashboard: create an account

| EduEval Manage |
| :--- |
| Brukerinfo |
| Bytte epost eller brukernavn |
| Username: |
| ibrahim |
| Email: |
| ibib@ntnu.no |
| Fullfor |
| Bytte passord |
| Current password: |
|  |
| New password: |
| Confirm new password: |
|  |
| Fullfor |

## 2) Add a course and add an exam



My Courses

Select a course to browse exams:
ID 101010

IE203612

## Exams for IE203612

Select exam to see and edit details:
Autumn_2018 ' ill

Add exam

## 3) Add number of questions and grade of each question

## Autumn_2018

| Back |
| :--- |
| Get evaluation link! |
| Number of que |
| Question 1: |
| 20 |
| Question 2: |
| 20 |
| Question 3: |
| 20 |
| Question 4: |
| 20 |
| Question 5: |
| 20 |
| Confirm |

Define default memberships

## Complexity

## 4) Define default membership functions



| very low | low | medium | high |
| :--- | :--- | :--- | :--- |

Importance


## 5) Get evaluation link (for students voting)

Autumn_2018


Confirm

Autumn_2018

## Back

Evaluation url:


Estimates by teacher
http://www.edueval.no/evaluate/W43VoKHPOnzIRXazLiP1dMIG

## a) Vote for complexity

## Evaluation of Autumn_2018



Define Complexity of questions

Question 1 very low low medium very high

## b) Vote for importance

Define Importance of questions


## c) Input time spent on each question

## Time spent on each question

```
Question 1:
Time spent (min)
```

```
Question 2:
```

Question 2:
Time spent (min)

```
    Time spent (min)
```

Question 3:
Time spent (min)

```
Question 4:
Time spent (min)
```


## Question 5:

```
Time spent (min)
```


## 6) Upload accuracy matrix of your students

## Current weights:

| Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Sum |
| :--- | :--- | :--- | :--- | :--- | :--- |
| 20 | 20 | 20 | 20 | 20 | 100 |

## Re-evaluated weights

| Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Sum |
| :--- | :--- | :--- | :--- | :--- | :--- |
| 19.79 | 19.27 | 21.89 | 19.79 | 19.27 | 100 |

Upload Accuracy Matrix


## Upload Accuracy Matrix

## Adjusted grades

Re-evaluated weights

| Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Sum |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 19.79 | 19.27 | 21.89 | 19.79 | 19.27 | 100 |
| Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Sum |
| 11.67 | 0.19 | 16.85 | 14.44 | 17.92 | 61.09 |
| 6.93 | 5.20 | 15.10 | 14.25 | 9.44 | 50.92 |
| 19.79 | 2.70 | 21.23 | 3.56 | 1.54 | 48.82 |
| 13.06 | 0.77 | 15.54 | 3.17 | 15.61 | 48.14 |
| 2.18 | 16.96 | 3.72 | 9.89 | 12.52 | 45.27 |
| 1.58 | 3.08 | 18.82 | 0.40 | 17.92 | 41.81 |
| 16.62 | 0.77 | 19.04 | 6.33 | 7.51 | 50.28 |
| 4.55 | 4.24 | 9.19 | 18.20 | 9.83 | 46.01 |
| 0.79 | 15.61 | 19.92 | 17.81 | 18.69 | 72.82 |
| 4.75 | 10.21 | 16.20 | 4.95 | 11.75 | 47.86 |

Upload Accuracy Matrix

## An example

- 10 students 5 questions

$$
\begin{aligned}
G^{T} & =\left[\begin{array}{lllll}
10 & 15 & 20 & 25 & 30
\end{array}\right] \\
I & =\left[\begin{array}{lllll}
0.9 & 0.434 & 0.87 & 0.1 & 0.486
\end{array}\right] \\
C & =\left[\begin{array}{lllll}
0.33 & 0.634 & 0.762 & 0.188 & 0.56
\end{array}\right],
\end{aligned}
$$


$A=\left[\begin{array}{cccccccccc}0.59 & 0.35 & 1 & 0.66 & 0.11 & 0.08 & 0.84 & 0.23 & 0.04 & 0.24 \\ 0.01 & 0.27 & 0.14 & 0.04 & 0.88 & 0.16 & 0.04 & 0.22 & 0.81 & 0.53 \\ 0.77 & 0.69 & 0.97 & 0.71 & 0.17 & 0.86 & 0.87 & 0.42 & 0.91 & 0.74 \\ 0.73 & 0.72 & 0.18 & 0.16 & 0.5 & 0.02 & 0.32 & 0.92 & 0.9 & 0.25 \\ 0.93 & 0.49 & 0.08 & 0.81 & 0.65 & 0.93 & 0.39 & 0.51 & 0.97 & 0.61\end{array}\right]$,
$T=\left[\begin{array}{cccccccccc}0.7 & 0.4 & 0.1 & 1 & 0.7 & 0.2 & 0.7 & 0.6 & 0.4 & 0.9 \\ 1 & 0 & 0.9 & 0.3 & 1 & 0.3 & 0.2 & 0.8 & 0 & 0.3 \\ 0 & 0.1 & 0 & 0.1 & 0.9 & 1 & 0.2 & 0.3 & 0.1 & 0.4 \\ 0.2 & 0.1 & 0 & 1 & 1 & 0.3 & 0.4 & 0.8 & 0.7 & 0.5 \\ 0 & 0.1 & 1 & 1 & 0.6 & 1 & 0.8 & 0.2 & 0.8 & 0.2\end{array}\right]$,

## Effective MFs after students' voting



Effective complexity


## Effective

 importance
## Comparison of three approaches

| Evaluation Method | Rank |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $1>$ | 2> | 3> | 4> | 5> | 6> | 7> | 8> | 9> | 10 |
| Classical | 9 | 1 | 2 | 8 | $10=$ | 4= | $5=$ | 6 | 7 | 3 |
| Fuzzy | 9 | 1 | 6 | 4 | 7 | 10 | 3 | 2 | 5 | 8 |
| Student involvement | 9 | 1 | 6 | 4 | 7 | 10 | 2 | 3 | 8 | 5 |

## Comparison of three approaches

| Evaluation <br> Method | Rank |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $1>$ | $2>$ | $3>$ | $4>$ | $5>$ | $6>$ | $7>$ | $8>$ | $9>$ | 10 |  |  |  |  |  |  |  |  |
| Classical | 9 | 1 | 2 | 8 | $10=$ | $4=$ | $5=$ | 6 | 7 | 3 |  |  |  |  |  |  |  |  |
| Fuzzy | 9 | 1 | 6 | 4 | 7 | 10 | 3 | 2 | 5 | 8 |  |  |  |  |  |  |  |  |
| Student <br> involvement | 9 | 1 | 6 | 4 | 7 | 10 | 2 | 3 | 8 | 5 |  |  |  |  |  |  |  |  |

- Fuzzy approaches are able to overcome the problem of ranking students of equal total scores.
- Student 3 ranked $10^{\text {th }}$ in classical approach becomes $7^{\text {th }}$ in fuzzy approach
- Student 2 ranked $3^{\text {rd }}$ in classical approach becomes $8^{\text {th }}$ in fuzzy approach
- Students 2 and 3 swapped ranks using students' involvement approach


## Discussion

- Student 3 has obtained better rates in questions 1 and 3 which are the most important questions in the exam (0.9 0.87)
- In addition, question 3 is the most difficult question in the exam with a difficulty ratio of 0.762 .
- When students are involved in the evaluation process resulted in a new MFs that considers most of the questions are more difficult than what the teacher was expecting and therefore student 3 lost the advantage of solving some of the most difficult questions.
- Feedback from students in the form of voting revealed a new fact that most of the exam question are to some degree difficult and very difficult.


## Future work

- Use of type 2 fuzzy sets to represent different views and more uncertainties can be handled.



## Acknowledgements

- The work presented in this paper was supported by the Center for Excellent IT Education (Exclted), project number 80428000 https://www.ntnu.edu/excited
- Stands for Centre for Excellent IT Education
- Exclted aims to enhance the learning in our study programs through increased student engagement and cross-campus collaboration and co-reflection among students and staff.
- The project start in Autumn 2017 and ended in Spring 2018
- Funds given: 50 KNOK
- Related courses: Intelligent systems IE303312


## TAKK for (meg)

## Questions

