Chapter 12:
Learning Analytics

Part 1: Learner Modelling
Part 2: Learner Adaptation

Pierre Dillenbourg and Patrick Jerman
1) Learner Modelling

\[ 5^2 = ? \]
Learner Modelling

From the learner’s behaviour, infer his/her learner’s knowledge state?
Learner Modelling

From the learner’s behaviour, infer his/her learner’s knowledge state

\[ 5^2 = 25 \quad \Rightarrow \quad \text{knows } X^2 \]

\[ 5^2 \neq 25 \quad \Rightarrow \quad \text{doesn’t know } X^2 \]
# Cognitive Diagnosis

<table>
<thead>
<tr>
<th>Behavior (Answer)</th>
<th>$5^2 = 25$</th>
<th>$5^n = \ldots$</th>
<th>$n^2 = n \cdot N$</th>
<th>$x^n = x \cdot x$ but bad mult.</th>
<th>$x^n = x \cdot n$</th>
<th>$x^n = x + n$</th>
<th>$x^n = ??$</th>
<th>Sum</th>
<th>Entropy</th>
<th>Normalized entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.10</td>
<td>0.20</td>
<td>0.30</td>
<td>0.40</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.89</td>
<td>0.63</td>
</tr>
<tr>
<td>35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.10</td>
<td>0.00</td>
<td>0.50</td>
<td>1</td>
<td>1.41</td>
<td>0.47</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.79</td>
<td>0.00</td>
<td>0.20</td>
<td>1</td>
<td>0.79</td>
<td>0.26</td>
</tr>
<tr>
<td>27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.00</td>
<td>0.00</td>
<td>0.59</td>
<td>1</td>
<td>1.03</td>
<td>0.34</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.59</td>
<td>0.40</td>
<td>1</td>
<td>1.03</td>
<td>0.34</td>
</tr>
</tbody>
</table>

**Diagnosis Power**: 0.41
Anxious / Self-confident
Risk-aversive / Risk-seeking
Aural / visual / kinesthetic
Deep / Surface
Field-dependent / independent

State ≠ Trait

Measured at time t  Stable in time

Severe criticisms:
• Contextual rather than personal
• No clear effects of adaptation
• Should education mimic style or counterbalance them?
• Labels produce self-fulfilling prophecies

Learning styles
Cognitive styles
BEWARE OF
the medicalisation of Education !!!

- Learning disabilities, LD
- Attention-deficit disorder, ADD
- Attention-deficit hyperactivity disorder, ADHD
- Dysgraphia, dyscalculia, dyslexia, ...
- High-potential children, HP
- ...

Labels help Sales
Learner Modelling

From the learner’s *behaviour*, infer his/her learner’s *knowledge state*

\[
\begin{align*}
    p (\text{state} = \text{knows} \mid \text{correct-answer}) & = 1 - \text{Guess} \\
    p (\text{state} = \text{knows} \mid \text{incorrect-answer}) & = 0 + \text{Slip}
\end{align*}
\]

*Factors that depend upon the response modality*

*Bayesian Knowledge Tracing, Corbezt & Anderson*
# Learner Modelling

## Knowledge States

<table>
<thead>
<tr>
<th>Behavior (Answer)</th>
<th>$5^2 = 25$</th>
<th>$5^n = \ldots$</th>
<th>$n^2 = n \cdot N$</th>
<th>$x^n = x \cdot x$ but bad mult.</th>
<th>Sum</th>
<th>Entropy</th>
<th>Normalized entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.10</td>
<td>0.20</td>
<td>0.30</td>
<td>0.40</td>
<td></td>
<td>1.89</td>
<td>0.63</td>
</tr>
<tr>
<td>35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td></td>
<td>1.41</td>
<td>0.47</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.79</td>
<td></td>
<td>0.79</td>
<td>0.26</td>
</tr>
<tr>
<td>27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td></td>
<td>1.03</td>
<td>0.34</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.59</td>
<td></td>
<td>1.03</td>
<td>0.34</td>
</tr>
</tbody>
</table>

## Knowledge States (with SLIP/GUESS factors)

<table>
<thead>
<tr>
<th>Behavior (Answer)</th>
<th>$5^2 = 25$</th>
<th>$5^n = \ldots$</th>
<th>$n^2 = n \cdot n$</th>
<th>$x^n = x \cdot x$ but bad mult.</th>
<th>Sum</th>
<th>Entropy</th>
<th>Normalized entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.05</td>
<td>0.15</td>
<td>0.25</td>
<td>0.35</td>
<td></td>
<td>2.53</td>
<td>0.84</td>
</tr>
<tr>
<td>35</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.34</td>
<td></td>
<td>2.16</td>
<td>0.72</td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.64</td>
<td></td>
<td>1.72</td>
<td>0.57</td>
</tr>
<tr>
<td>27</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.54</td>
<td></td>
<td>1.87</td>
<td>0.62</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.50</td>
<td></td>
<td>1.92</td>
<td>0.64</td>
</tr>
</tbody>
</table>
From the learner’s behaviours, infer his/her learner’s knowledge state

\[ b(s) = \text{watch video with many pauses} \]

\[ b(s) = \text{select correct definition of SD in a quiz with 5 possible definitions} \]

\[ b(s) = \text{post a message “There is a mistake on the slide” (and there is one indeed)} \]

State “fine”: the learner is performing well
State “active”: the learner is working but does not seem to succeed well
State “lost”: the learner does not understand at all or did not complete the activities
State “drop”: the learned has dropped out (e.g. no login since N days)

Normalized entropy of the diagnosis vector

\[ X(S) = \{\text{lost, active, fine, brilliant}\} \]

\[ x(s) = [0.15, 0.40, 0.30, 0.15] \quad H_0 = 0.94 \]

\[ x(s) = [0.05, 0.15, 0.25, 0.55] \quad H_0 = 0.80 \]

\[ x(s) = [0.01, 0.02, 0.02, 0.95] \quad H_0 = 0.18 \]
The uncertainty of the diagnosis can be estimated by Shannon’s entropy applied to the vector for probabilities for the different states.

Since this value depends upon the number of states, we normalize it on a 0→1 scale by dividing it by the maximal entropy which \( \log_2 \) of the number of states.

The diagnosis power of a question can be measured by the entropy of the diagnosis vector.

\[
H(X) = - \sum_i P(x_i) \log_b P(x_i)
\]
Which question has the highest diagnosis power?

Question 1

• The standard deviation of a distribution if the \( \text{sum of } f \text{ from the mean} \)

Question 2

• Remove two numbers from this distribution to minimize it’s standard deviation : [1 3 3 5 9 9 9 10 11 18 19 25 29]
Basic approach to reduce uncertainty

Decrease uncertainty by collecting multiple answers

<table>
<thead>
<tr>
<th>$5^2 = ??$</th>
<th>Knowledge States (with SLIP/GUESS factors)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Sum</th>
<th>Entropy</th>
<th>Normalized entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior (Answer)</td>
<td>$5^2 = 25$</td>
<td>$5^n = ...$</td>
<td>$n^2 = n.n$</td>
<td>$x^0 = x.x ...$</td>
<td>$x^0 = x . x$ but bad mult.</td>
<td>$x^n = x . n$</td>
<td>$x^n = x + n$</td>
<td>$x^n = ??$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.05</td>
<td>0.15</td>
<td>0.25</td>
<td>0.35</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>1</td>
<td>2.53</td>
</tr>
<tr>
<td>35</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.34</td>
<td>0.05</td>
<td>0.00</td>
<td>0.41</td>
<td>1</td>
<td>2.16</td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.00</td>
<td>0.64</td>
<td>0.00</td>
<td>0.15</td>
<td>1</td>
<td>1.72</td>
</tr>
<tr>
<td>27</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.54</td>
<td>1</td>
<td>1.87</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.50</td>
<td>0.30</td>
<td>1</td>
<td>1.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$7^2 = ??$</th>
<th>Knowledge States (second question)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Sum</th>
<th>Entropy</th>
<th>Normalized entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior (Answer)</td>
<td>$7^2 = 25$</td>
<td>$7^n = ...$</td>
<td>$n^2 = n.n$</td>
<td>$x^0 = x.x ...$</td>
<td>$x^0 = x . x$ but bad mult.</td>
<td>$x^n = x . n$</td>
<td>$x^n = x + n$</td>
<td>$x^n = ??$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>0.02</td>
<td>0.10</td>
<td>0.35</td>
<td>0.45</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>1</td>
<td>1.95</td>
</tr>
<tr>
<td>56</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.61</td>
<td>0.00</td>
<td>0.00</td>
<td>0.31</td>
<td>1</td>
<td>1.43</td>
</tr>
<tr>
<td>14</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.82</td>
<td>0.00</td>
<td>0.10</td>
<td>1</td>
<td>1.04</td>
</tr>
<tr>
<td>72</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.32</td>
<td>0.00</td>
<td>0.00</td>
<td>0.60</td>
<td>1</td>
<td>1.44</td>
</tr>
<tr>
<td>9</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.72</td>
<td>0.20</td>
<td>1</td>
<td>1.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$3^3 = ??$</th>
<th>Knowledge States (second question)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Sum</th>
<th>Entropy</th>
<th>Normalized entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior (Answer)</td>
<td>$3^3 = 25$</td>
<td>$3^n = ...$</td>
<td>$n^3 = n.n.n$</td>
<td>$x^0 = x.x ...$</td>
<td>$x^0 = x . x$ but bad mult.</td>
<td>$x^n = x . n$</td>
<td>$x^n = x + n$</td>
<td>$x^n = ??$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>0.01</td>
<td>0.10</td>
<td>0.20</td>
<td>0.65</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>1</td>
<td>1.53</td>
</tr>
<tr>
<td>26</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.81</td>
<td>0.00</td>
<td>0.00</td>
<td>0.15</td>
<td>1</td>
<td>0.94</td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.82</td>
<td>0.00</td>
<td>0.14</td>
<td>1</td>
<td>0.92</td>
</tr>
<tr>
<td>33</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.12</td>
<td>0.00</td>
<td>0.00</td>
<td>0.84</td>
<td>1</td>
<td>0.86</td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.91</td>
<td>0.05</td>
<td>1</td>
<td>0.63</td>
</tr>
</tbody>
</table>
How does the teacher/system chooses the next question?

• Because it will maximize the learning gain of the learner?

   Exploitation

• Because it will maximize the system knowledge about the learner?

   Exploration
Exploration Exploitation Tradeoff

Learner 1
Learner 2
Learner 3
Learner 4
Learner 5
Learner 6
Learner 7
Learner 8
Learner 9
Learner Modelling

From the learner’s behaviour, **infer** his/her learner’s knowledge state

\[ \triangle \text{(learner-knowledge, correct-knowledge)} = f \left( \triangle \text{(learner-answer, correct-answer)} \right) \]

L1: 3 cm

L2: 8 cm

Correct answer = 8.54
Learner Answer = 8.18 = SQRT (8^2 + 3^1)
Inference mechanisms: If, when bringing perturbation X to an expert system, it produces the same mistake as the learner, X is a good hypothesis of what the learner did not understand.

From the learner’s behaviour, infer his/her learner’s knowledge state.

Correct Knowledge → Incorrect Knowledge → Learner’s Knowledge

- Correct Response
- Remove a rule
- Add a ‘malrule’
- Hypothesis
- Learner’s Errors
From the learner’s behaviour, infer his/her learner’s knowledge state.
From the learner’s behaviour, infer his/her learner’s knowledge state

Behaviours

- Text entered
- Item selected (button, menu,…)
- Area clicked
- Line drawn with mouse.
- Response time
- …
- …
- Number of pauses
- Mouse path
- Gaze path
- Facial expressions
- Gestures
- …

Behavioural ‘Dust’

(fragments of behaviour that do not have an explicit semantic value)

Classifier

State A

Inference mechanisms

State B
From the learner’s behaviour, infer his/her learner’s knowledge state

<table>
<thead>
<tr>
<th>Plane</th>
<th>Behaviours</th>
<th>Behavioural ‘Dust’</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Class Plane</td>
<td>The # messages in the forum</td>
<td>Head Co-Rotation</td>
</tr>
<tr>
<td>2 Team Plane</td>
<td>The concept map produced by a pair</td>
<td>Gaze Recurrence</td>
</tr>
<tr>
<td>1 Individual Plane</td>
<td>The learner answer to a quizz</td>
<td>Video ‘Withmeness’</td>
</tr>
</tbody>
</table>
From the learner’s behaviour, infer his/her learner’s knowledge state

(Beyond the sake of cognitive research), it is only interesting to discriminate state X and Y, if the next decision will be different for X and Y.
From the learner’s behaviour and his previous state, infer his/her learner’s knowledge state.
BKT & Orchestration Graphs

Inferring the learner’s state from his previous state
Dropped Out

Active

Lost

Fine

The weight of edges
## Orchestration graph: State Transition Matrix

### Activity 2

<table>
<thead>
<tr>
<th></th>
<th>Lost</th>
<th>Active</th>
<th>Fine</th>
<th>Great</th>
<th>$H$</th>
<th>$H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost</td>
<td>0.26</td>
<td>0.39</td>
<td>0.29</td>
<td>0.06</td>
<td>1.80</td>
<td>0.90</td>
</tr>
<tr>
<td>Active</td>
<td>0.19</td>
<td>0.34</td>
<td>0.26</td>
<td>0.21</td>
<td>1.96</td>
<td>0.98</td>
</tr>
<tr>
<td>Fine</td>
<td>0.11</td>
<td>0.28</td>
<td>0.45</td>
<td>0.16</td>
<td>1.81</td>
<td>0.90</td>
</tr>
<tr>
<td>Great</td>
<td>0.05</td>
<td>0.15</td>
<td>0.25</td>
<td>0.55</td>
<td>1.60</td>
<td>0.80</td>
</tr>
</tbody>
</table>

$1 - H_0 = 0.10$
<table>
<thead>
<tr>
<th></th>
<th>M6</th>
<th>M7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lost</td>
<td>Active</td>
</tr>
<tr>
<td>Lost</td>
<td>0.01</td>
<td>0.24</td>
</tr>
<tr>
<td>Active</td>
<td>0.01</td>
<td>0.24</td>
</tr>
<tr>
<td>Fine</td>
<td>0.01</td>
<td>0.24</td>
</tr>
</tbody>
</table>

ω(M5) 0.45

ω(M6) 0.45

Same predictability but opposite predictions
### State Transition Matrix

\[
\begin{align*}
\text{M8} & \begin{bmatrix}
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
\end{bmatrix} \\
\text{M9} & \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\end{bmatrix} \\
\text{M10} & \begin{bmatrix}
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
\end{bmatrix} \\
\text{M11} & \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \\
\text{M12} & \begin{bmatrix}
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.1 & 0.1 & 0.2 & 0.3 & 0.3 \\
0 & 0 & 0.2 & 0.3 & 0.5 \\
0 & 0.1 & 0.2 & 0.2 & 0.4 \\
0 & 0 & 0.2 & 0 & 0.8 \\
\end{bmatrix} \\
\text{M13} & \begin{bmatrix}
0.5 & 0.1 & 0.2 & 0.1 & 0.1 \\
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.7 & 0.2 & 0.1 & 0 & 0 \\
0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\
0.8 & 0.2 & 0 & 0 & 0 \\
\end{bmatrix}
\end{align*}
\]

\[
\begin{align*}
\text{U(M)} & = 0 \\
\text{U(M)} & = 0.47 \\
\text{U(M)} & = 1 \\
\text{U(M)} & = -1 \\
\text{U(M)} & = 0.47 \\
\text{U(M)} & = -0.42
\end{align*}
\]
Learning Analytics: probabilities VS visualisations
Learner Modelling in Orchestration Graphs:

Third Dimension: one may infer the learner’s state from his behaviour (depth), his previous state (horizontally) and the state of others (vertically)
The learning analytics cube: 3 axes of inference

A. John does probably not understand SD deviation because he removed the central values of the distribution

B. John does probably not understand SD deviation because he did not understand what is a mean and the mean is a pre-requisite

C. John does probably not understand SD deviation because most learners in that class failed and John is one of the weakest
This cube may allow second-level inferences.
This cube may allow second-level inferences.
This cube may allow second-level inferences.
Machine Learning for Modelling Learning

Behaviours

Text entered
Item selected (button, menu, …)
Area clicked
Line drawn with mouse.
Response time
…
…
Number of pauses
Mouse path
Gaze path
Facial expressions
Gestures
…

Behavioural ‘Dust’

(fragments of behaviour that do not have an explicit semantic value)

State A

State B

Classifier

Inference mechanisms
Machine Learning for Modelling Learning

Row behavioural traces

Behavioural features

Classifier

State A

State B
Relevant Behavioral Abstractions (Features)

Computational Models → Education Research
gaze(a) = f(gaze(b))
DUET - Dual Eye-Tracking Pair programming experiment

Low gaze recurrence

P. Jermann, M.-A. Nüssli & P. Dillenbourg
© CRAFT – http://craft.epfl.ch/

Supported by the Swiss National Science Foundation (grants #K-12K1-117909 and #PZ00P_126611)
DUET - Dual Eye-Tracking Pair programming experiment

High gaze recurrence

P. Jermann, M.-A. Nüssli & P. Dillenbourg
© CRAFT – http://craft.epfl.ch/

Supported by the Swiss National Science Foundation (grants #K-12K1-117909 and #PZ00P_126611)
gaze(listener) = \( f(\text{gaze(speaker)}) \)

**Feature:** Gaze recurrence  
**Context:** Collaborative learning
Eye tracking experiment on MOOC Video

Following teacher’s references

Gaze of students’ watching Scala course by Prof. Martin Odersky (EPFL, Switzerland)

K. Sharma, P. Jermann, P. Dillenbourg
@ CHILI – http://chili.epfl.ch
Supported by the Swiss National Science Foundation
(Grants CR1211_132996 and PZ00P2_126611)
Relevant Behavioral Abstractions

gaze (learner) = \( f \) (reference (teacher))

**Feature:** Withmeness  
**Context:** Lecturing
“...they look like a bunch of little grains arranged together...typically a group of very small elements”
Do finger-based or gaze-based deictics enhance learning?

Sarah d’Angelo, Kshitij Sharma, Darren Gergle, Pierre Dillenbourg (2016)
Relevant Behavioral Abstractions

gaze (learner) = f (gaze (teacher))

Feature: ‘Withmeness’
Context: Lecturing
Modeling in the wild?

<table>
<thead>
<tr>
<th>Kernel</th>
<th>Features</th>
<th>Score</th>
<th>Cohen's kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF(c=1.31, g=0.0211)</td>
<td>Distance, Head travel norm., Num. still periods</td>
<td>61.86%</td>
<td>0.30</td>
</tr>
<tr>
<td>RF(c=1.21, g=0.11)</td>
<td>Period, Row, Head travel norm., Mean duration still</td>
<td>61.72%</td>
<td>0.32</td>
</tr>
<tr>
<td>RBF(c=1.11, g=0.061)</td>
<td>Head travel norm., Mean duration still</td>
<td>60.42%</td>
<td>0.28</td>
</tr>
<tr>
<td>RBF(c=1.4, g=0.04)</td>
<td>Period, Distance, Row, Mean duration still</td>
<td>59.23%</td>
<td>0.30</td>
</tr>
</tbody>
</table>
Relevant Behavioral Abstractions

gaze (learner) = f (location (teacher))

Feature: Head rotations
Context: Lecturing
activity (teacher) = \( f \) (gaze (teacher))
Relevant Behavioral Abstractions

activity(teacher) = \( f \) (location (teacher))

Feature: pupil diameter, \#faces in field of view
Context: Lecturing
Education raises challenges to data science

• Explainability: which features make sense
• Exploration/Exploitation trade-off
• Cold Start: simulations, expert’s knowledge
Education is a computational science

EPFL Center for Learning Sciences