

Interactive machine learning: a constructivist interpretation

Advait Sarkar



UNIVERSITY OF
CAMBRIDGE
Computer Laboratory



Constructivist learning theory offers fresh perspectives and design implications for interactive machine learning.

What is constructivism?

Constructivism is the view that human learning is the *construction of knowledge* as a result of the interaction between a person's ideas and their experiences, and is not merely the transfer of information.

What constructivist issues relate to IML?

1. **Task ownership:** a relevant, owned problem is intrinsic motivation for learning.

→ In *BrainCel*, users build models on their own data in spreadsheets.

2. **Ill-defined problem:** requires users to tackle emergent issues and make defensible judgements.

→ Evaluating the fit and quality of a statistical model is an ill-defined process.

3. **Perturbation:** nonconforming experiences necessitate mental 'accomodation'.

→ Errors in prediction cause users to re-evaluate the model.

4. **Reflexivity:** self-critical learners understand their thoughts and knowledge better.

→ Visualisations show how the model evolved over time in response to user actions.

New issues which remain unaddressed:

5. **Collaboration:** the construction of meaning is a social activity.

6. **Task in context:** historical and socio-cultural influences shape learning.

7. **Tool mediation:** systems are responsible for changing the culture which creates them.

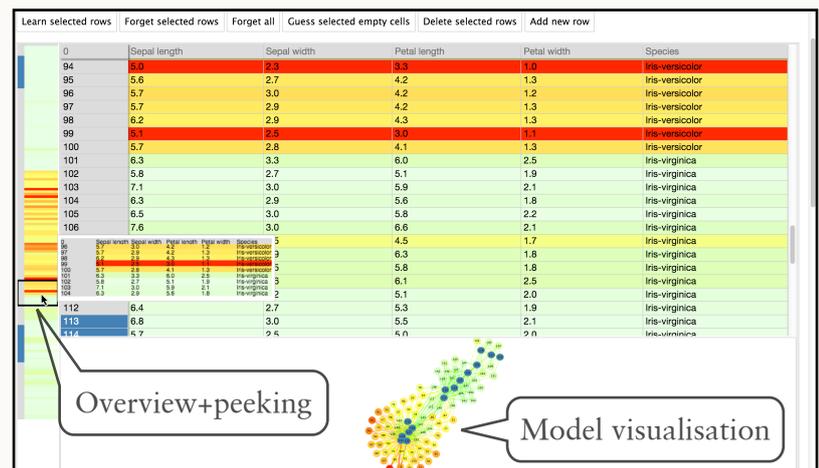
What is interactive machine learning?

Interactive machine learning (IML) lets non-experts build and apply machine learning models for their own use in a variety of scenarios. Typically, to construct a model, the user repeatedly chooses training data examples and assigns labels to them.

Example: Interactive machine learning in spreadsheets

In *BrainCel*, we augmented the spreadsheet, a versatile yet familiar data manipulation tool, to enable non-expert end-users to build and apply machine learning models.

The user selects rows of complete, correct data to 'teach' the software, which implicitly trains a statistical model. The user can then apply this model to guess values for empty cells, or validate the contents of filled cells.



Part of the *BrainCel* interface. Some data has been loaded into the spreadsheet. Rows with blue row numbers have been added by the user to the training set. Rows are coloured according to the model's confidence: green for high confidence, red for low.

Further reading

Sarkar, A. (2016). Constructivist Design for Interactive Machine Learning. In proceedings of the 34th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA 2016)

Sarkar, A., Jannik, M., Blackwell, A. F., & Spott, M. (2015). Interactive visual machine learning in spreadsheets. In 2015 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC) (pp. 159–163).