HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

User Modelling for Avoiding Overfitting in Interactive Knowledge Elicitation for Prediction

IUI 2018, Tokyo, Japan
08.03.2018

Pedram Daee, Tomi Peltola, Aki Vehtari, Samuel Kaski

Aalto University

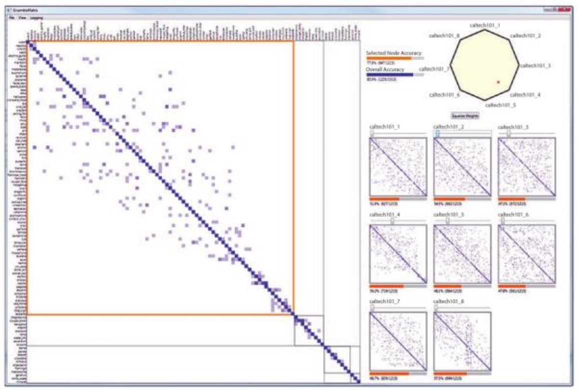
UNIVERSITY OF HELSINKI

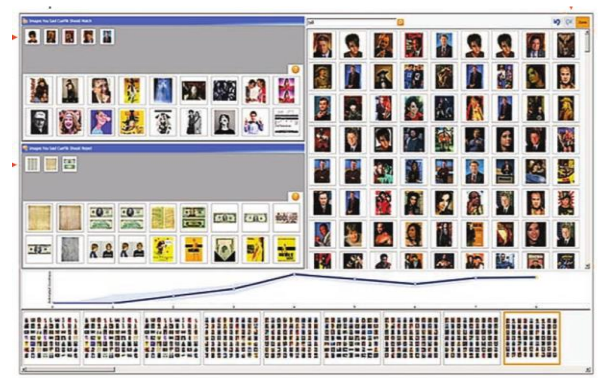
Outline

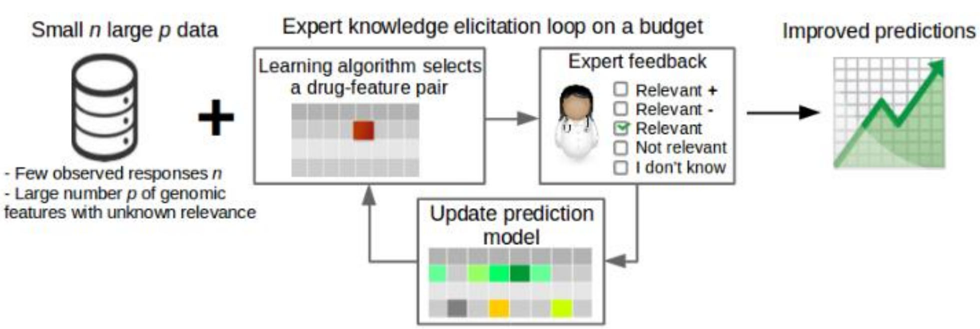
1. Introduction
   - Overfitting in IML? Is it important?
2. User modeling
   - Let the machine know the human biases
3. User study
   - Human machine collaboration for prediction
4. Summary

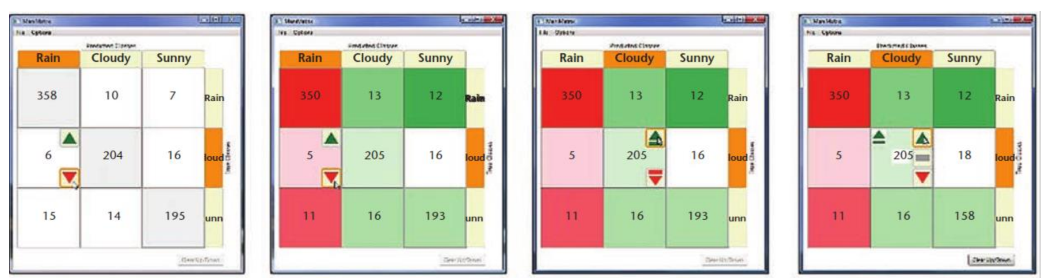
Introduction - Motivation

1. Humans are being added to the machine learning loop.









[Amershi, et.al. 2014]
[Sundin, et.al. 2017]

Introduction - Motivation

1. Humans are being added to the machine learning loop.
2. Humans can be biased too heavily on the piece of information offered.
   - Example:

   Students were given anchors that were obviously wrong. They were first asked whether Mahatma Gandhi died before or after age 9, or before or after age 140.

   Q: How old was Gandhi when he died?

   A: average age of 50 vs. average age of 67

[Tversky 1974]
[Garthwaite 2005]

Introduction - Motivation

1. Humans are being added to the machine learning loop.
2. Humans can be biased too heavily on the piece of information offered.

In interactive machine learning, machines need to account for potential human biases

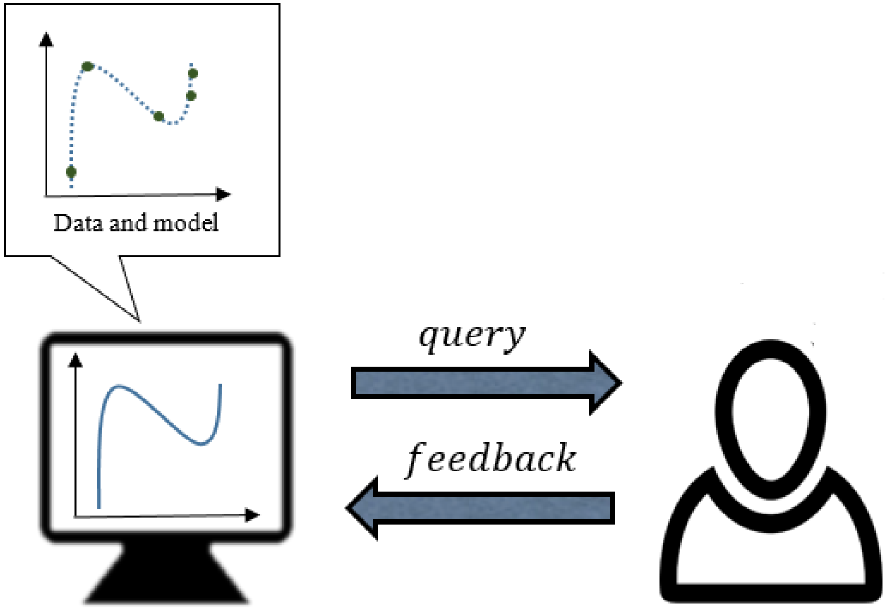
Introduction - Conceptual Scenario

- A human and a machine collaborate to solve a prediction problem.
- The machine has access to some data, but external knowledge is **required** to improve the results.
  - E.g. small *n* large *p* and other ill-posed problems
- The human may overfit to the training data through human-computer interactions
  - directly seeing the training data
  - through interaction

Introduction - Conceptual Scenario

Machine and **Human** are collaborating and machine queries the user

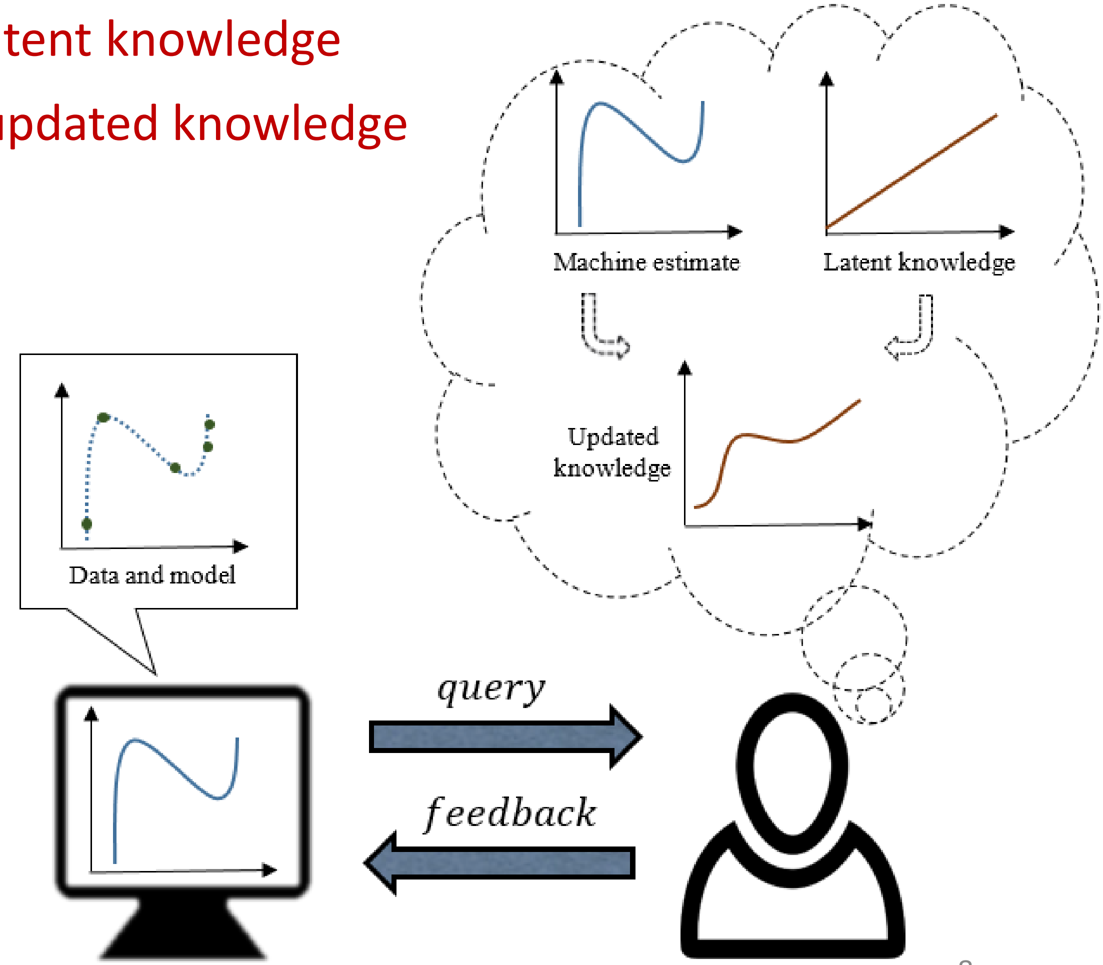
1. Fit a regression, visualize for the user



Introduction - Conceptual Scenario

Machine and **Human** are collaborating and machine queries the user

1. Fit a regression, visualize for the user
2. Observe, combine with latent knowledge
3. Give feedback based on updated knowledge

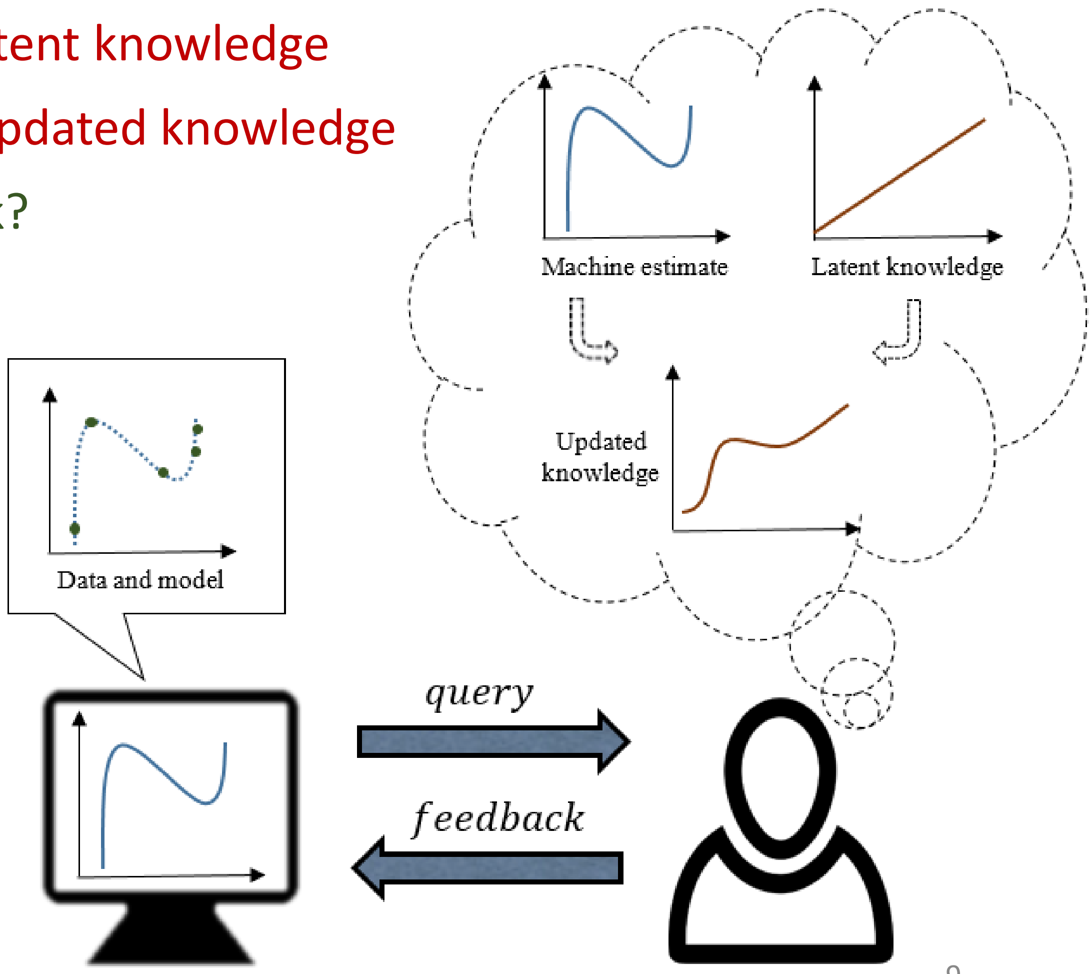


Introduction - Conceptual Scenario

Machine and **Human** are collaborating and machine queries the user

1. Fit a regression, visualize for the user
2. Observe, combine with latent knowledge
3. Give feedback based on updated knowledge
4. Directly use that feedback?

   Account for the bias first?

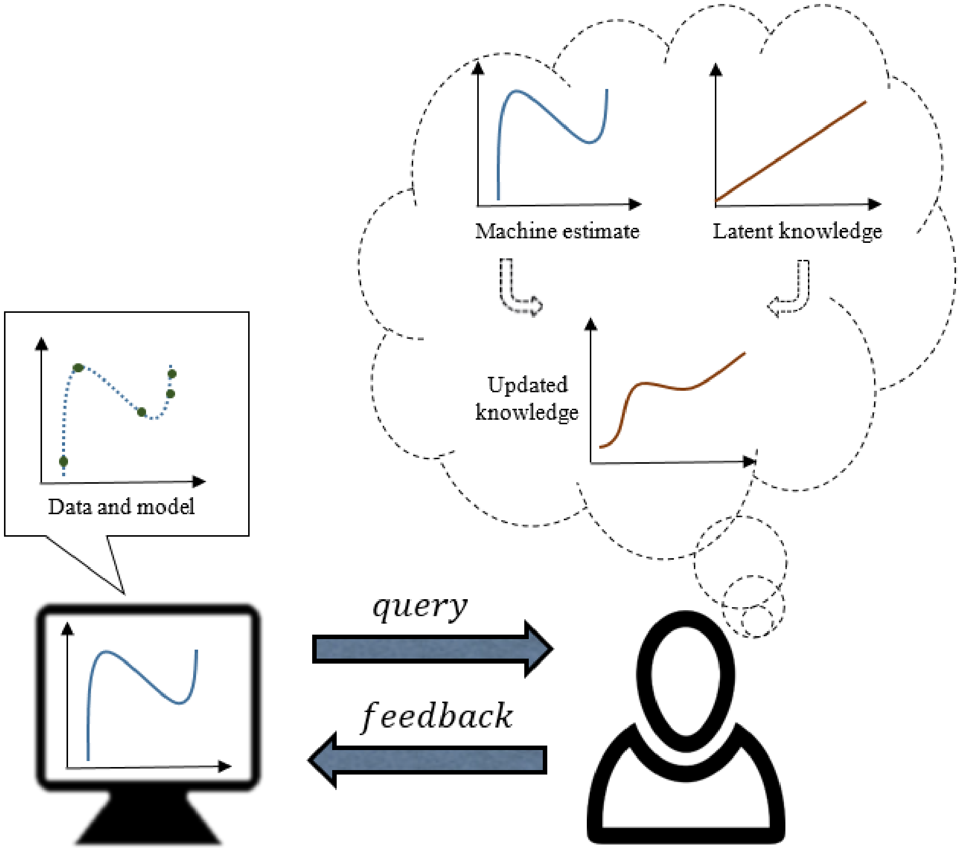


Introduction - Contribution

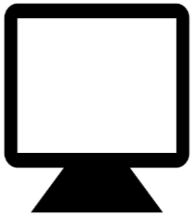
- If the human information is in part based on the training data, there would be a risk for double use of data and overfitting.
- We propose a user modelling methodology, by assuming simple rational behaviour, to correct this problem.
  - If the model is correct, then it may be able to undo the overfitting

User modeling - Undoing the bias

- SOLUTION:
  - The user latent knowledge can be inferred, by performing the inverse of what happens in the user mind
- TOOL:
  - Bayesian inference



User modeling - Undoing the bias



Bayesian rational user!

Training data D

Prior $P(\theta)$

Posterior $P(\theta|D) \propto P(\theta)P(D|\theta)$

$\hat{P}(\theta|D)$

User knowledge K with latent likelihood $P(K|\theta)$

$\hat{P}(\theta|D,K) \propto \hat{P}(\theta|D)P(K|\theta)$

$\hat{P}(\theta|D,K)$

Inferred user likelihood:

$$\hat{P}(K|\theta) \propto \frac{\hat{P}(\theta|D,K)}{\hat{P}(\theta|D)}$$

Final Posterior:
$P(\theta|D,K) \propto P(\theta)P(D|\theta)\,\hat{P}(K|\theta)$

User Study - Task

- Main task: Predict Amazon review ratings from its text (bag of words).

Cast Iron Is Awesome!

By RayJoni07 on March 12, 2015

Size: 12-Inch | Verified Purchase

This is my first time cooking with cast iron and from day one I've use this hefty skillet everyday! I love it! It's heavy but it's a work horse and it's so much better than my non stick pans I'm replacing! Eggs cook beautifully in this baby! I've fried fish, did homemade pancakes, bacon, turkey bacon, ground turkey!!! They all were perfection! The more it's used the better it really does get. It may take a little longer to get hot but it holds heat very well! Clean up is easy. Wash in hot water, dry, put on stove then give light oil wipe! And soap is ok as long as the season is good. I did an extra seasoning for an hour as recommended but have been making sure I use it often to get more coats on! It's really not as taxing as it seems! I've made our entire family a full breakfast from this one pan! The silicone sleeve is a great tool as well. The price is worth it! So cast iron for my nonstick and stainless is next for the rest of the set I need to replace. Buy this if you are considering because it's wonderful! The size is perfect for my family of 6! My next task this week will be deep dish pizza!

User Study - Task

- Main task: Predict Amazon review ratings from its text (bag of words).

Cast Iron Is Awesome!
By RayJoni07 on March 12, 2015

Size: 12-Inch | Verified Purchase

This is my first time cooking with cast iron and from day one I've use this hefty skillet everyday! I love it! It's heavy but it's a work horse and it's so much better than my non stick pans I'm replacing! Eggs cook beautifully in this baby! I've fried fish, did homemade pancakes, bacon, turkey bacon, ground turkey!!! They all were perfection! The more it's used the better it really does get. It may take a little longer to get hot but it holds heat very well! Clean up is easy. Wash in hot water, dry, put on stove then give light oil wipe! And soap is ok as long as the season is good. I did an extra seasoning for an hour as recommended but have been making sure I use it often to get more coats on! It's really not as taxing as it seems! I've made our entire family a full breakfast from this one pan! The silicone sleeve is a great tool as well. The price is worth it! So cast iron for my nonstick and stainless is next for the rest of the set I need to replace. Buy this if you are considering because it's wonderful! The size is perfect for my family of 6! My next task this week will be deep dish pizza!

“is” and “a” may not be informative for score prediction

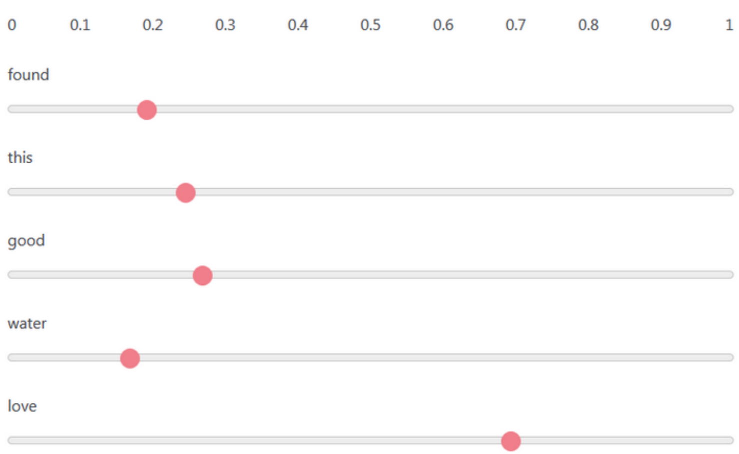
“perfect” and “love” indicate high review score

User Study - Task

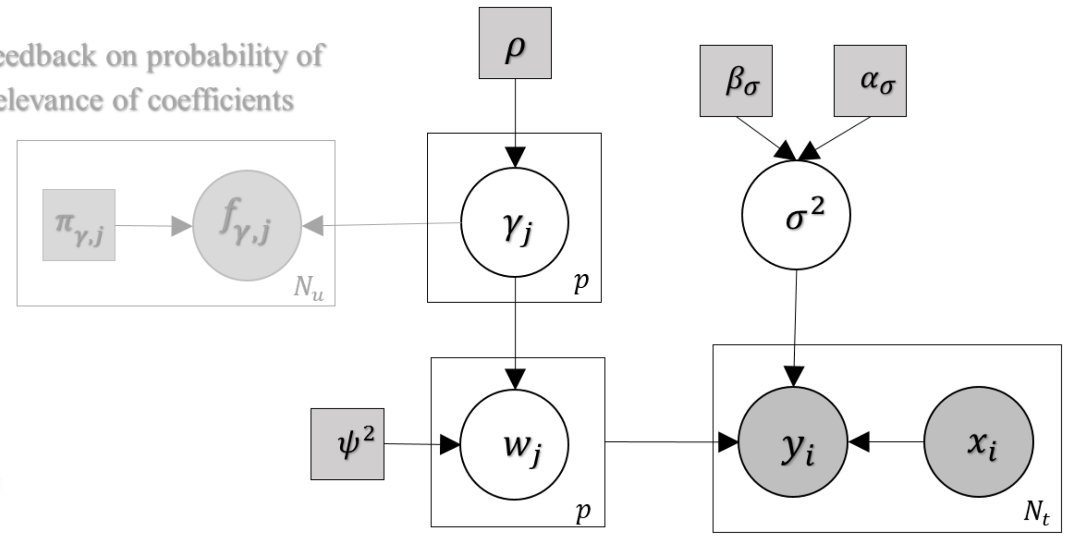
- Main task: Predict Amazon review ratings from its text (bag of words).
- User's task:
  - Estimate the relevance of a set of words for the prediction (scale 0-1; UI: sliders).

User Study - Task

- Main task: Predict Amazon review ratings from its text (bag of words).
- User's task:
  - Estimate the relevance of a set of words for the prediction (scale 0-1; UI: sliders).
- Conditions for user's feedback:
  - Baseline: sliders set to default position.
  - Interactive elicitation (IE): sliders set to machine's estimate of relevance.



User Study - Model for collaboration



$$
\begin{aligned}
\boldsymbol{y} &\sim \mathrm{N}(\boldsymbol{X}\boldsymbol{w}, \sigma^2\,\mathbf{I}),\\
\sigma^{-2} &\sim \mathrm{Gamma}(\alpha_\sigma, \beta_\sigma),\\
w_j &\sim \gamma_j\,\mathrm{N}(0, \psi^2) + (1-\gamma_j)\delta_0,\\
\gamma_j &\sim \mathrm{Ber}(\rho),\\
f_{\gamma,j} &\sim \gamma_j\,\mathrm{Ber}(\pi_{\gamma,j}) + (1-\gamma_j)\,\mathrm{Ber}(1-\pi_{\gamma,j})
\end{aligned}
$$

[Daee, et.al. 2017]

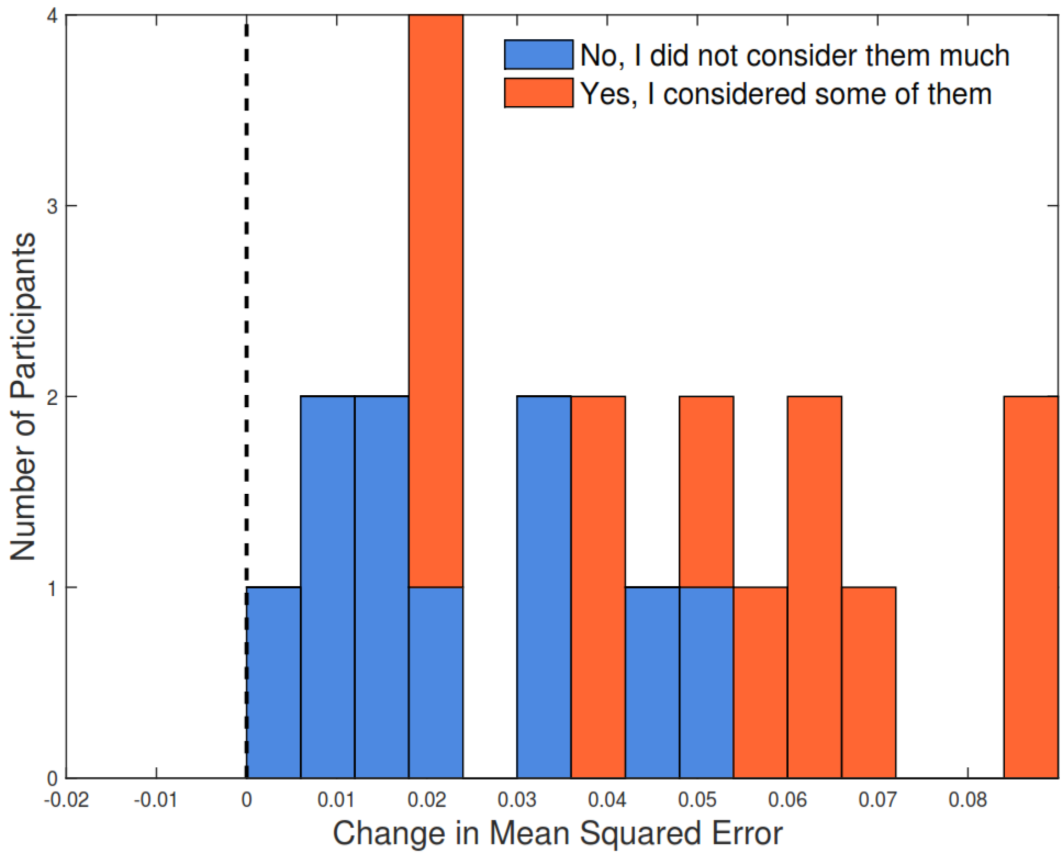
User Study - Results (1/3)

- 48 participants, 3 excluded.
  - 23 in baseline, 22 in IE system
- Mean Squared Error (MSE) on test data.

	Test MSE $\pm$ STD		
Condition	*No feedback*	*User feedback*	*User model*
Baseline	1.835	1.749 ± 0.050	NA
IE	1.835	1.744 ± 0.045	1.705 ± 0.038

User Study - Results (2/3)

- For IE system, how much the user model improved the test error?
  - For all users, the user model improved the MSE



User Study - Results (3/3)

- Was the feedback between the two groups different?

Keyword	*Machine*	*IE*	*Baseline*	*P-value*
best	0.89	0.90	0.80	0.014
great	1.00	0.95	0.83	0.031
disappointed	0.99	0.96	0.85	0.038
heavy	0.20	0.36	0.52	0.038
buy this	0.70	0.77	0.63	0.053
steel	0.19	0.12	0.24	0.073
line	0.34	0.15	0.06	0.088
don't	0.47	0.54	0.38	0.090
recommend	0.16	0.74	0.84	0.102
good	0.26	0.71	0.81	0.115

Summary

- The traditional assumption, that the human just neutrally and passively gives feedback in IML, may not hold in practice.
- We described a user modelling methodology for disentangling latent user knowledge from observed user feedback that was given in response to machine revealing information from the training data.
- The IML results can be improved by user modeling.
- Codes and data:
  - https://github.com/HIIT/human-overfitting-in-IML

References

- Daee, P., Peltola, T., Soare, M., Kaski, S. (2017). Knowledge elicitation via sequential probabilistic inference for high-dimensional prediction. *Machine Learning.*
- Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*
- Sundin, et al. (2017). Improving drug sensitivity predictions in precision medicine through active expert knowledge elicitation. *arXiv preprint arXiv:1705.03290*.
- Garthwaite, P. H., Kadane, J. B., & O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*
- Tversky, A., Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science.

Thank you!

HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY