LLM Processes

Numerical Predictive Distributions Conditioned on Natural Language

James Requeima
University of Toronto

John Bronskill
University of Cambridge

Dami Choi
University of Toronto

Rich Turner
University of Cambridge

David Duvenaud
University of Toronto
Vector Institute

We'd Like to Incorporate Textual Information when Modelling

- Textual information is currently underutilized in probabilistic regression:
  - Doctors notes, news articles, weather reports, csv headers…
- Incorporating prior information is hard.

LLM-Powered Probabilistic Prediction

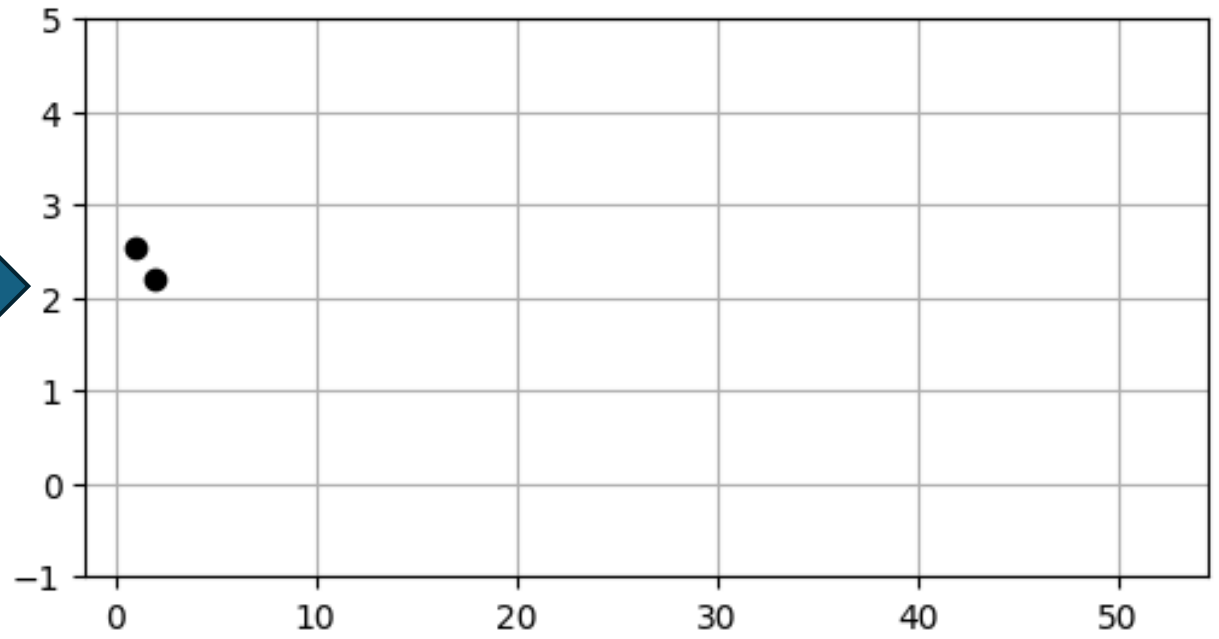
LLM Processes allow you to make predictions:

1. By directly explaining, in plain language, the unique information that you have about a particular problem.
2. That harness the massive latent knowledge in SOTA LLMs.

We developed a Regression Model That You Can Talk To.

User provides information to the model in plain language.

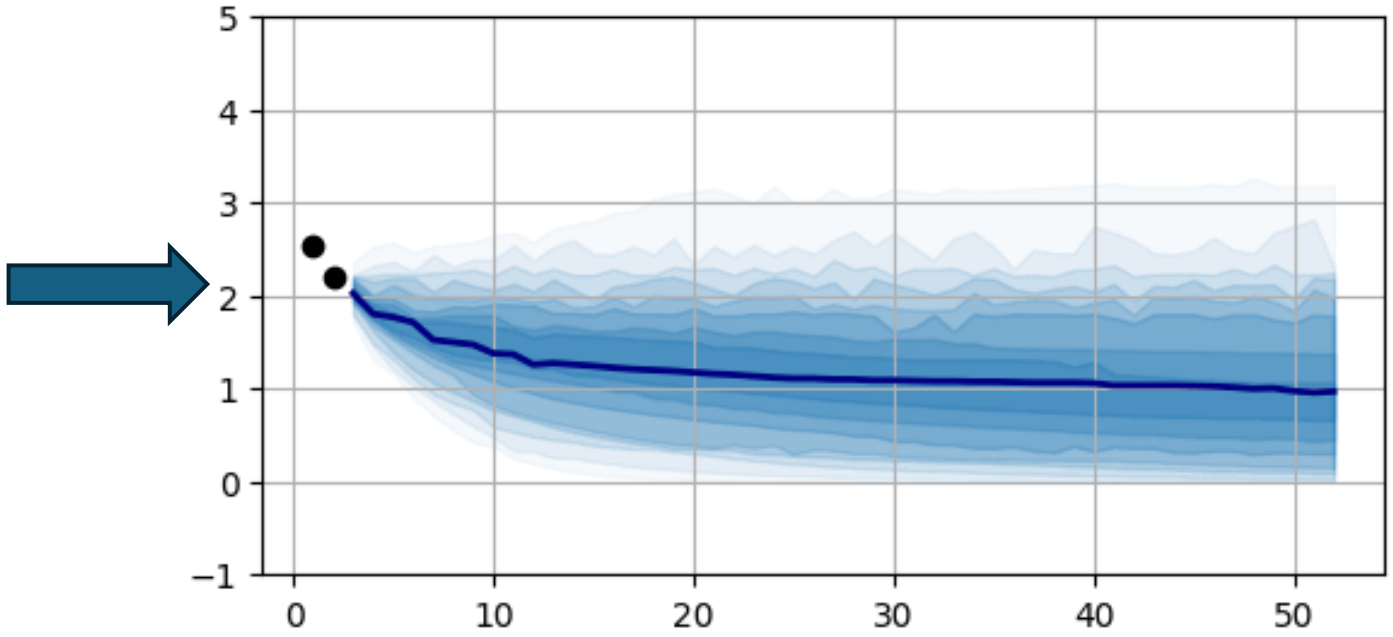
Training points shown to the model



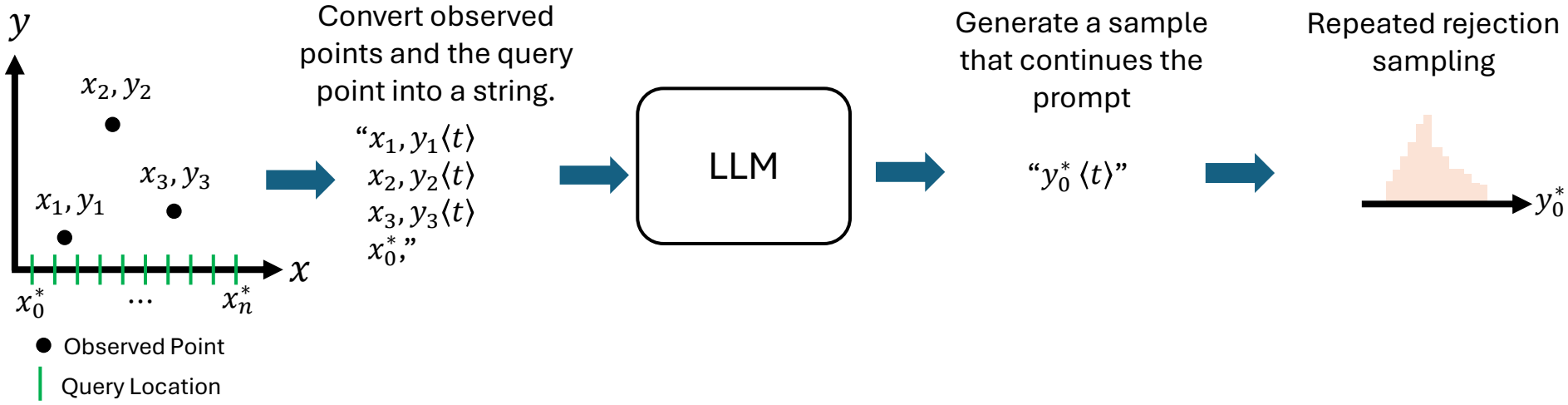
We developed a Regression Model That You Can Talk To.

User provides information to the model in plain language.

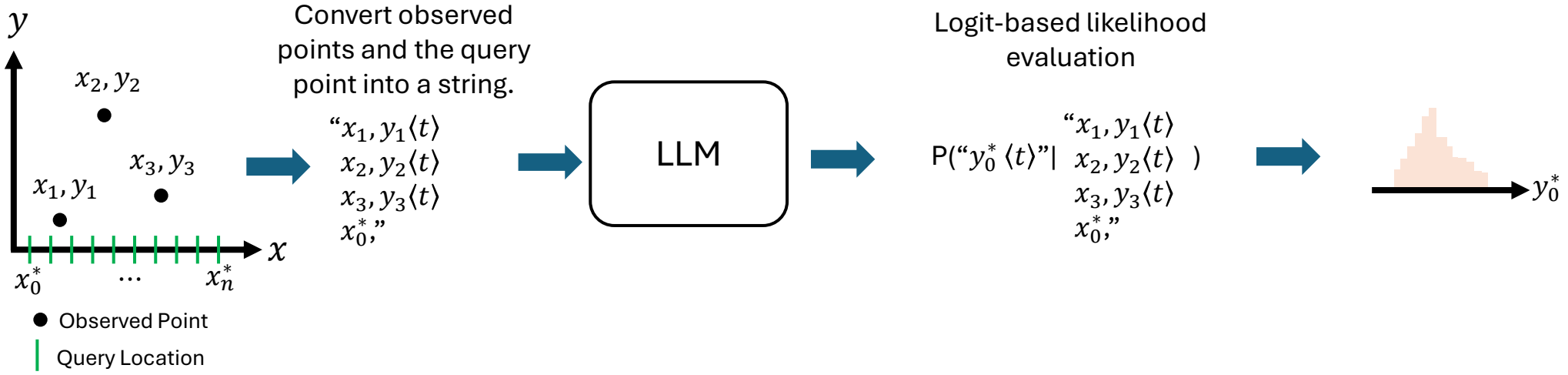
Training points shown to the model



Eliciting a Predictive Distribution from an LLM

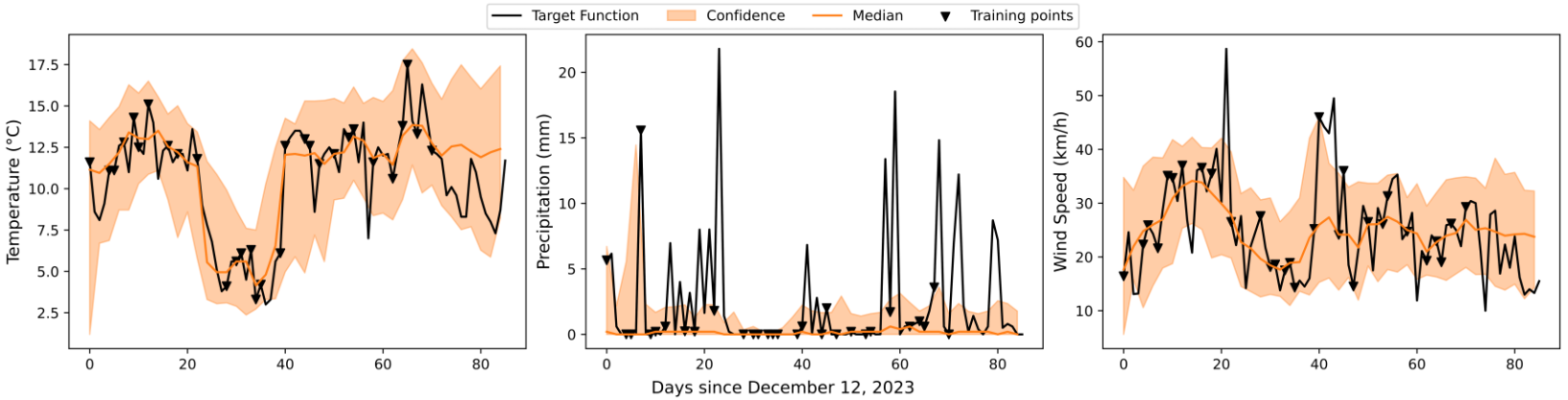


Eliciting a Predictive Distribution from an LLM



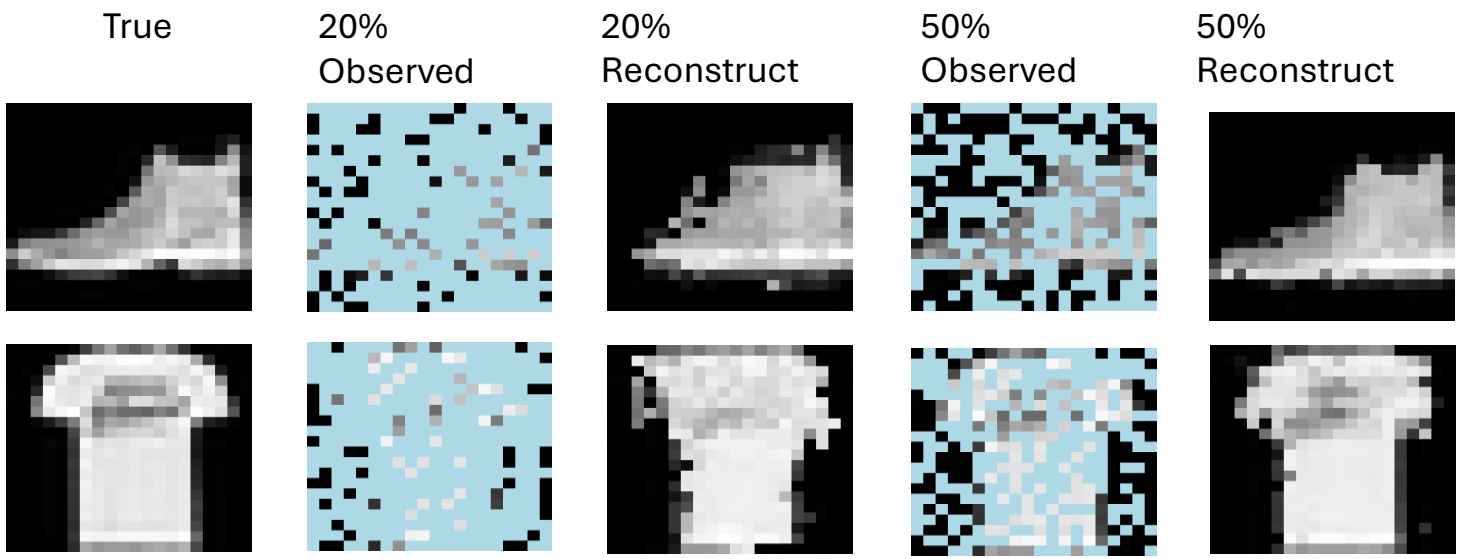
LLMPs Have Flexible Data Handling

Multidimensional data: weather prediction (1D in, 3D out)



Llama-3-7B LLMP

Image Reconstruction (2D in, 1D out)



Mixtral 8×7B LLMP

Incorporating Side Information from Text

Task: Predict US housing prices using various features given 10 example houses.

- Case 1: model is given features as only numerical values

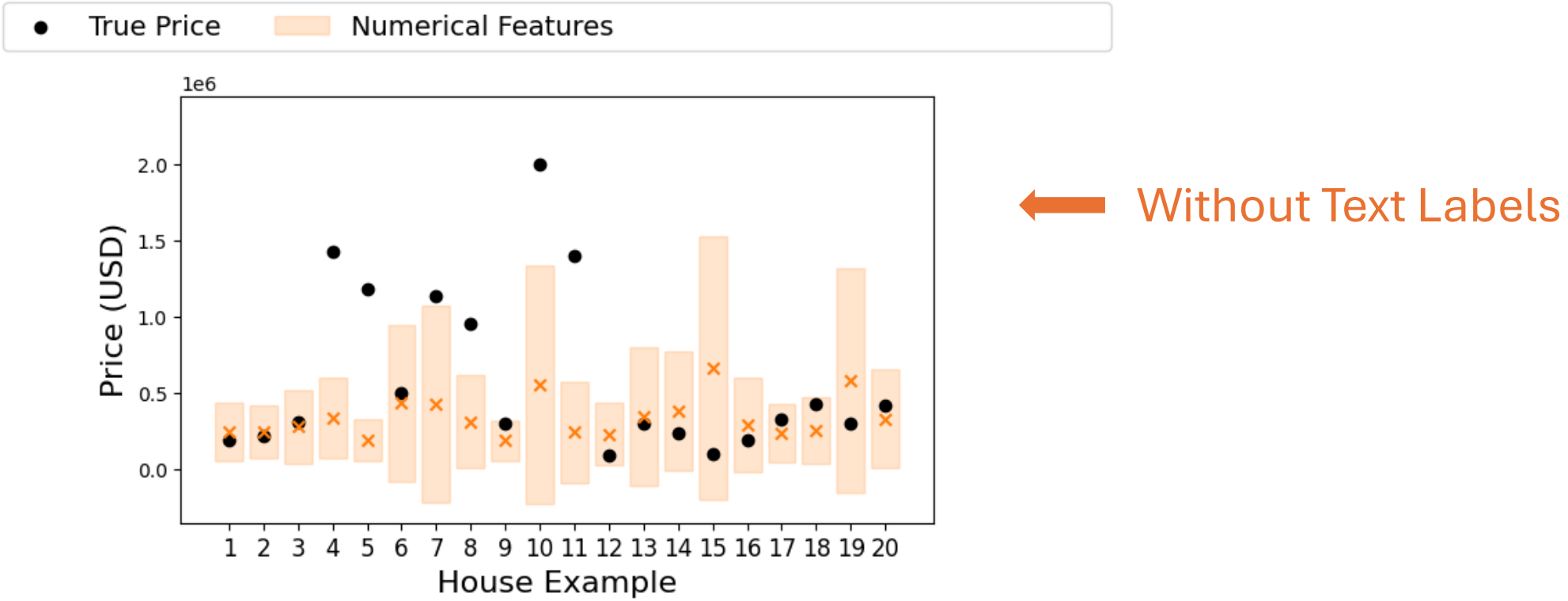
"30.45738, -97.75516, 78729, 107830.0, 30907, 1216.1, 1349, 3"

- Case 2: model is given numerical features with text labels.

"Location: Austin, Texas, Latitude: 30.45738, Longitude: -97.75516, Zip Code: 78729, Median Household Income: 107830.0, Zip Code Population: 30907 people, Zip Code Density: 1216.1, people per square mile, Living Space: 1349 square feet, Number of Bedrooms: 3, Number of, Bathrooms: 2"

Same numerical information.

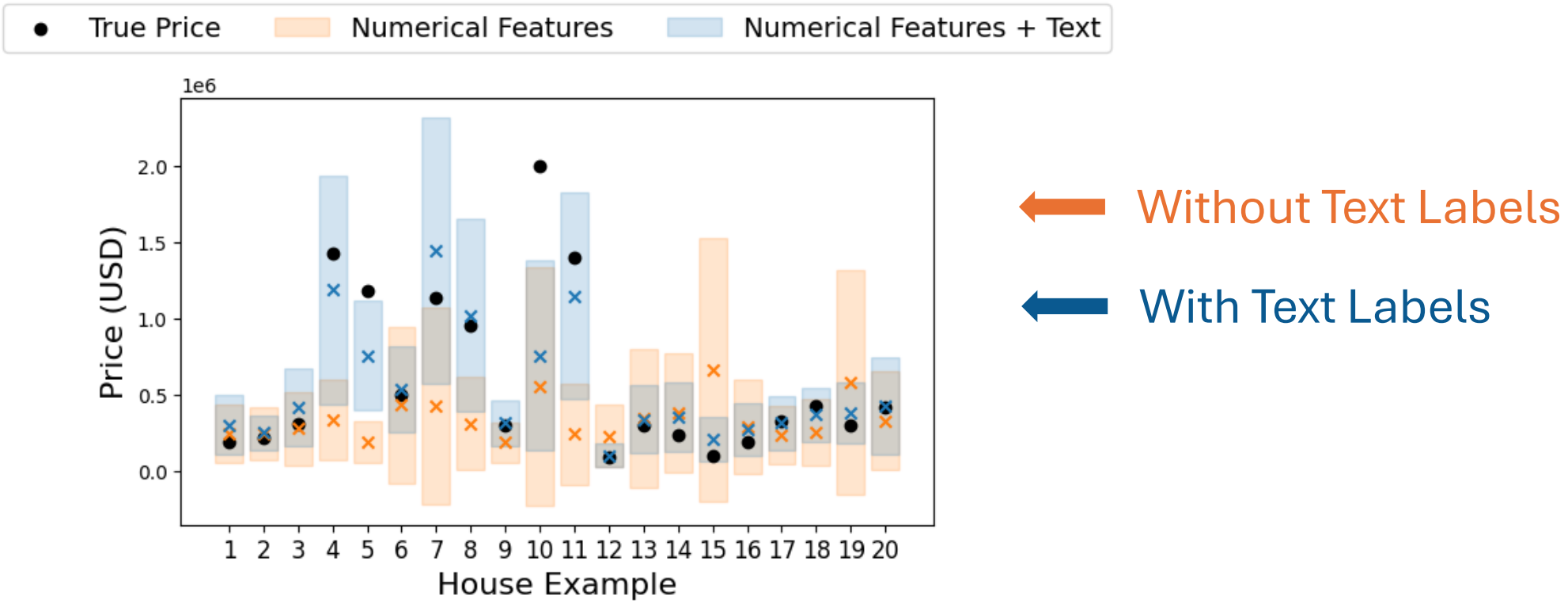
Incorporating Side Information from Text



← Without Text Labels

Using Mixtral-8×7B Instruction Tuned

Incorporating Side Information from Text

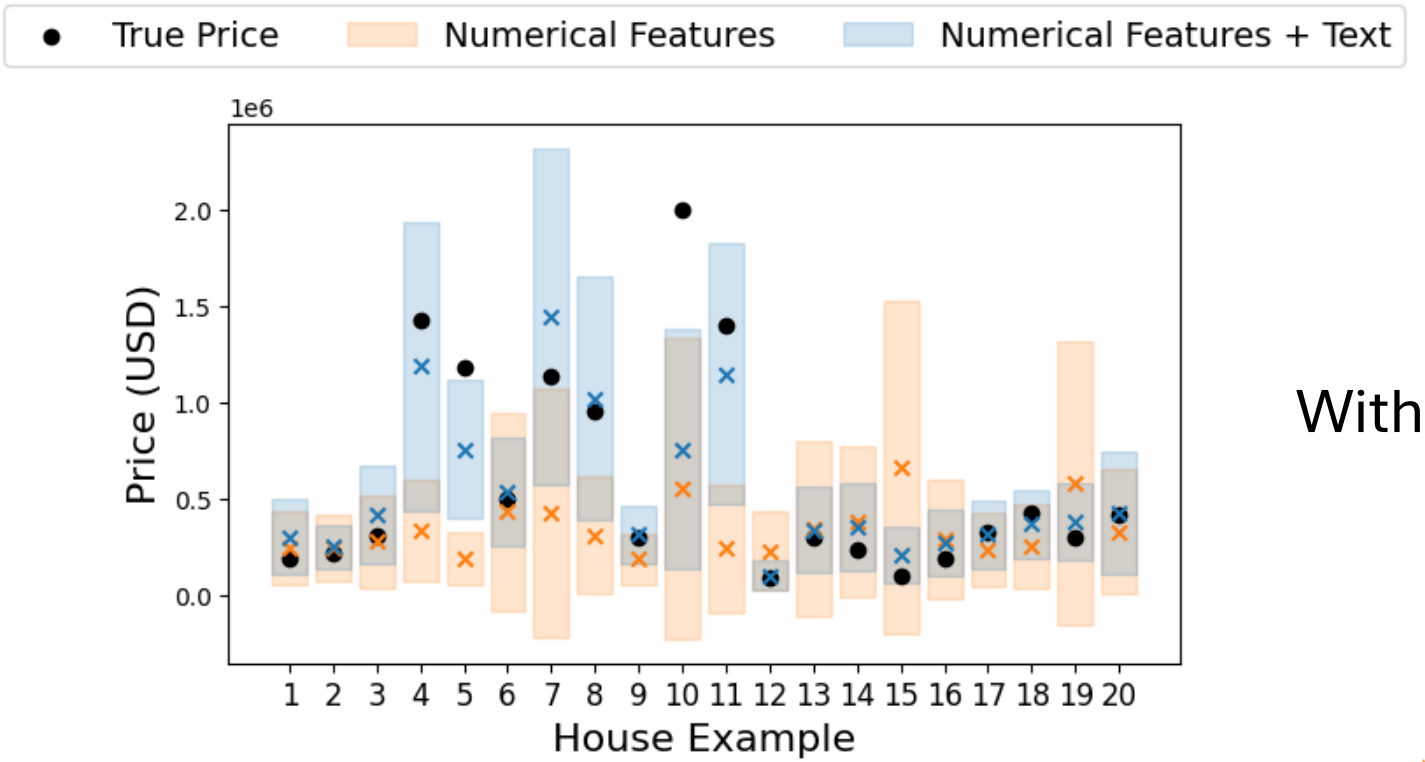


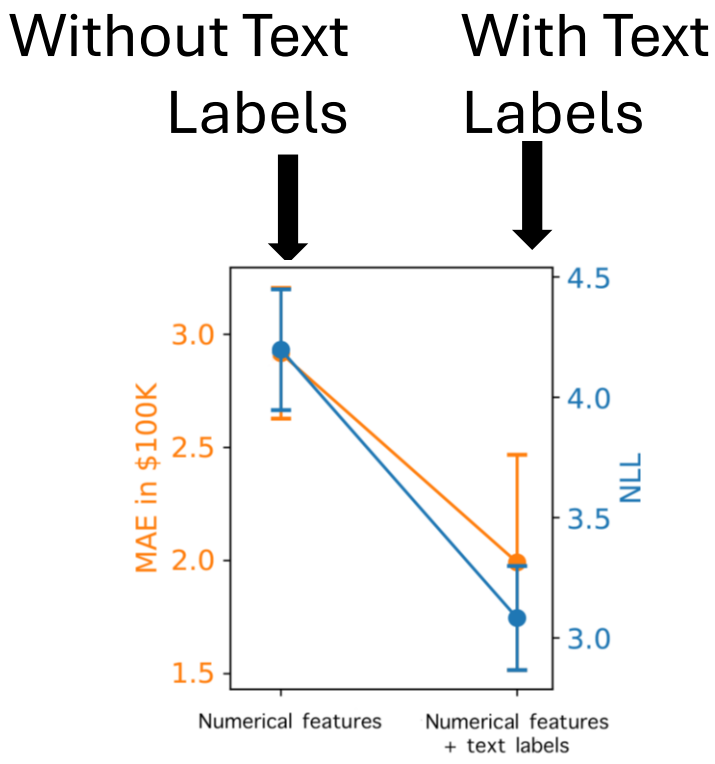
Without Text Labels

With Text Labels

Using Mixtral-8×7B Instruction Tuned

Incorporating Side Information from Text





Using Mixtral-8×7B Instruction Tuned

Thank You!

Paper

Code

James Requeima
University of Toronto

John Bronskill
University of Cambridge

Dami Choi
University of Toronto

Rich Turner
University of Cambridge

David Duvenaud
University of Toronto
Vector Institute