PROBABILISTIC MACHINE LEARNING FOR OCCUPANCY PREDICTION BASED ON SENSOR FUSION

Amirreza Heidari
amirreza.heidari@epfl.ch

Dolaana Kholayg
dolaana.khovalyg@epfl.ch

Verena marie barthelmes
verena.barthelmes@epfl.ch

Thermal Engineering for the Built Environment Laboratory
EPFL, Lausanne, Switzerland

Why to predict occupancy?

All the building services are to satisfy the occupants

- Automatic schedule of air conditioning
- Adapt ventilation rates
- Regulate temperature set points
- Manage workspace

A possible approach is to use camera

But it is:

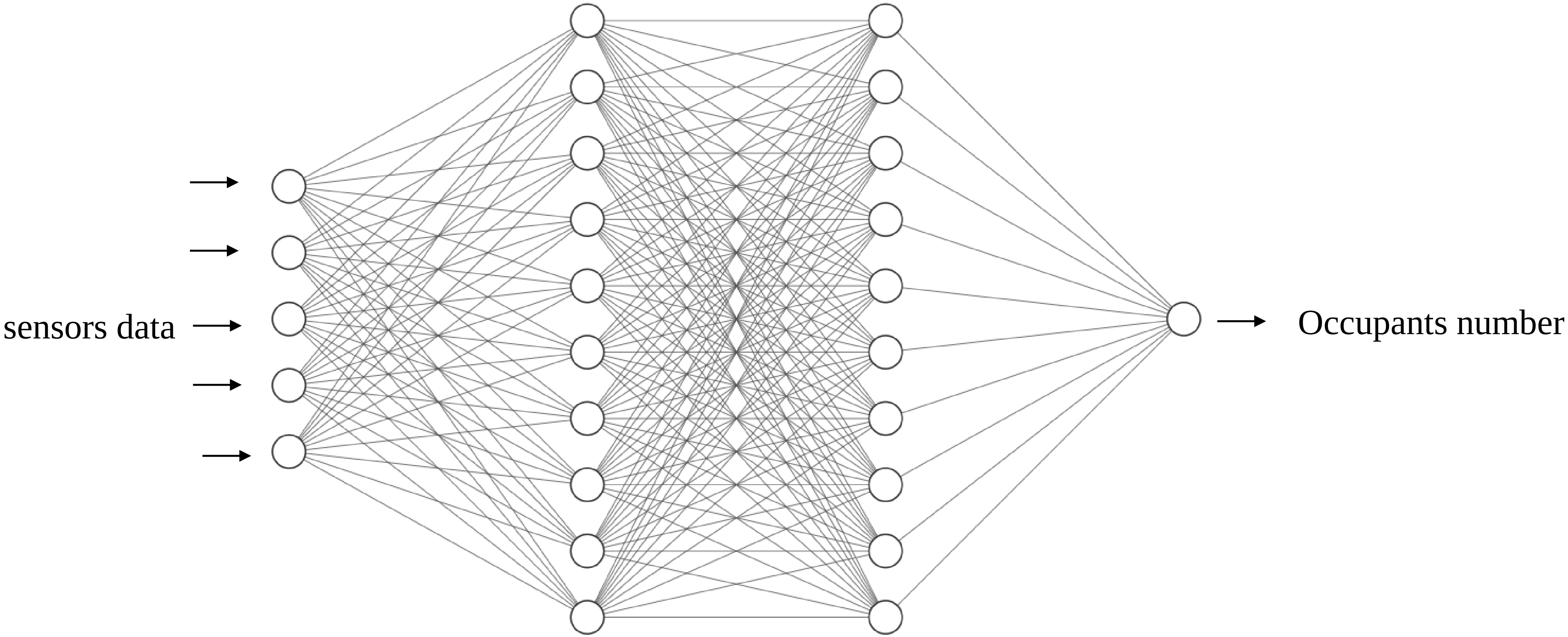
- Expensive
- Intrusive
- Complicated to process and automate

An alternative:
Implicit occupancy sensing based on sensor fusion

Use a combination of sensors that are expected to be correlated with occupancy

- Cheap sensors
- Commonly exist in some buildings
- Non-intrusive

Machine Learning for occupancy prediction



Limitation of conventional methods

Human behavior is highly stochastic

Conventional models are deterministic:

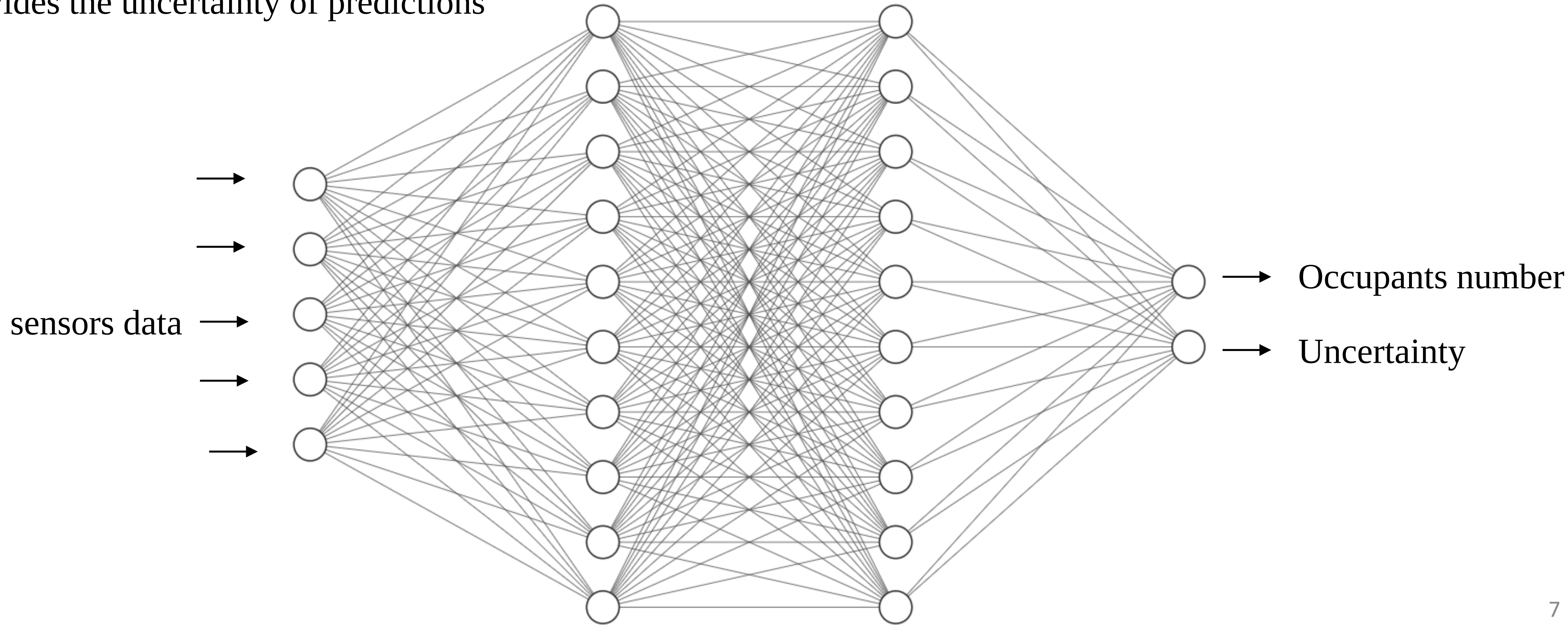
- predict a value confidently
- Do not provide any information about how model is certain about the predicted number

While having a good accuracy, exact number of occupants is not predicted most of the time

Research aim

A probabilistic deep learning model which:

- Predicts number of occupants
- Provides the uncertainty of predictions



Methodology

Monte-Carlo Dropout presented by Gal and Ghahramani on 2016 [1]

Probabilistic loss function

$$-log\ \varphi_{\theta}(x) = \frac{log\ \hat{\sigma}_{\theta}{}^{2}(x)}{2} - \frac{(y - \hat{\mu}_{\theta}(x))^{2}}{2\hat{\sigma}_{\theta}{}^{2}(x)}$$

Activation of Monte-Carlo Dropout during test phase

$$\bar{\mu}(x) = \frac{1}{N}\sum_{n \epsilon N} \bar{\mu}_{n}(x)$$

$$\bar{\sigma}^{2}(x) = \frac{1}{N}\sum_{n \epsilon N} (\hat{\sigma}_{\theta}{}^{2}(x) + \bar{\mu}_{n}{}^{2}(x)) - y(x)^{2}$$

Case study

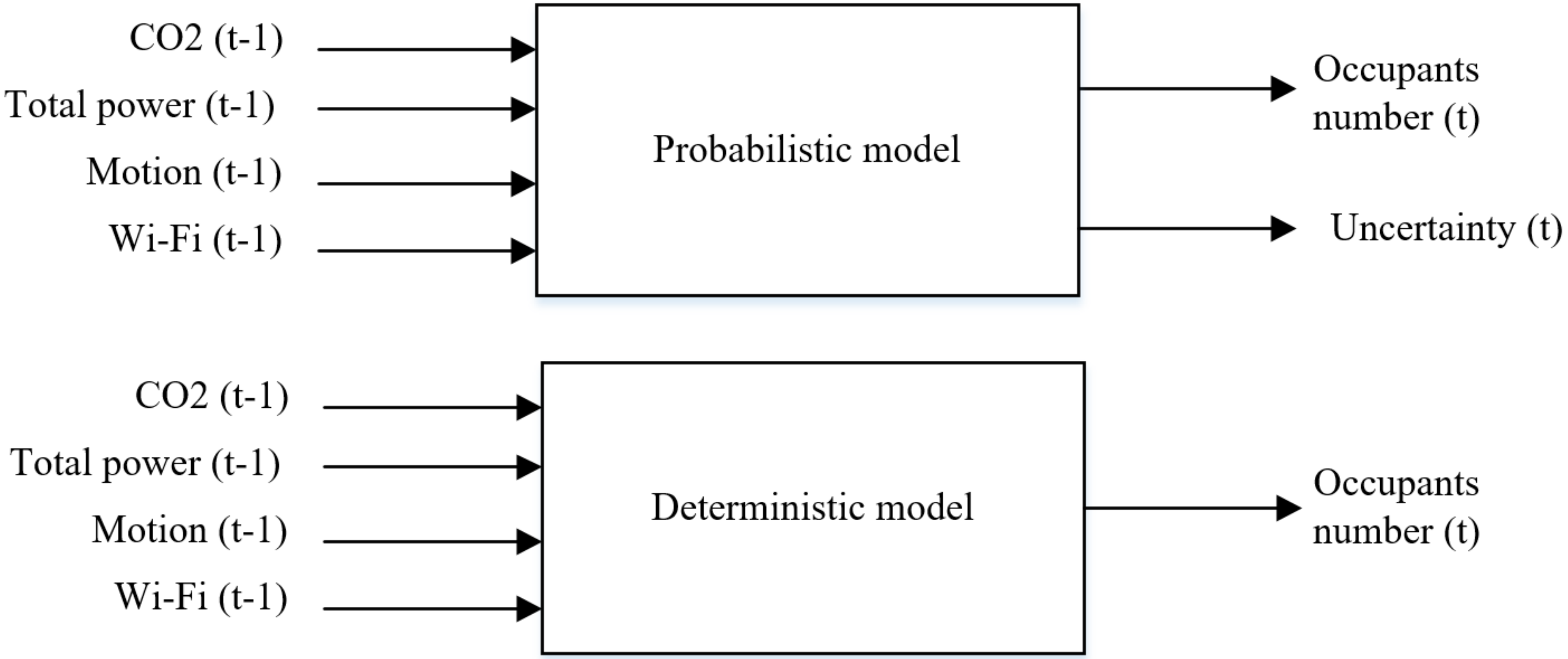
- A 4 floor academic office building in Ottawa, Canada
- measurements include:
  - ✓ CO_2 concentration
  - ✓ Detected motions
  - ✓ Plug and lighting power use
  - ✓ Total power use
  - ✓ Number of connected WiFi devices
- Ground truth data provided by camera on the entrance and exit points
- Public dataset published by Hobson et al. [2]

Case study

Location of different sensors in floors 3 and 4 (plans edited from [2])

Developed models

- Two probabilistic models: LSTM, Feed Forward



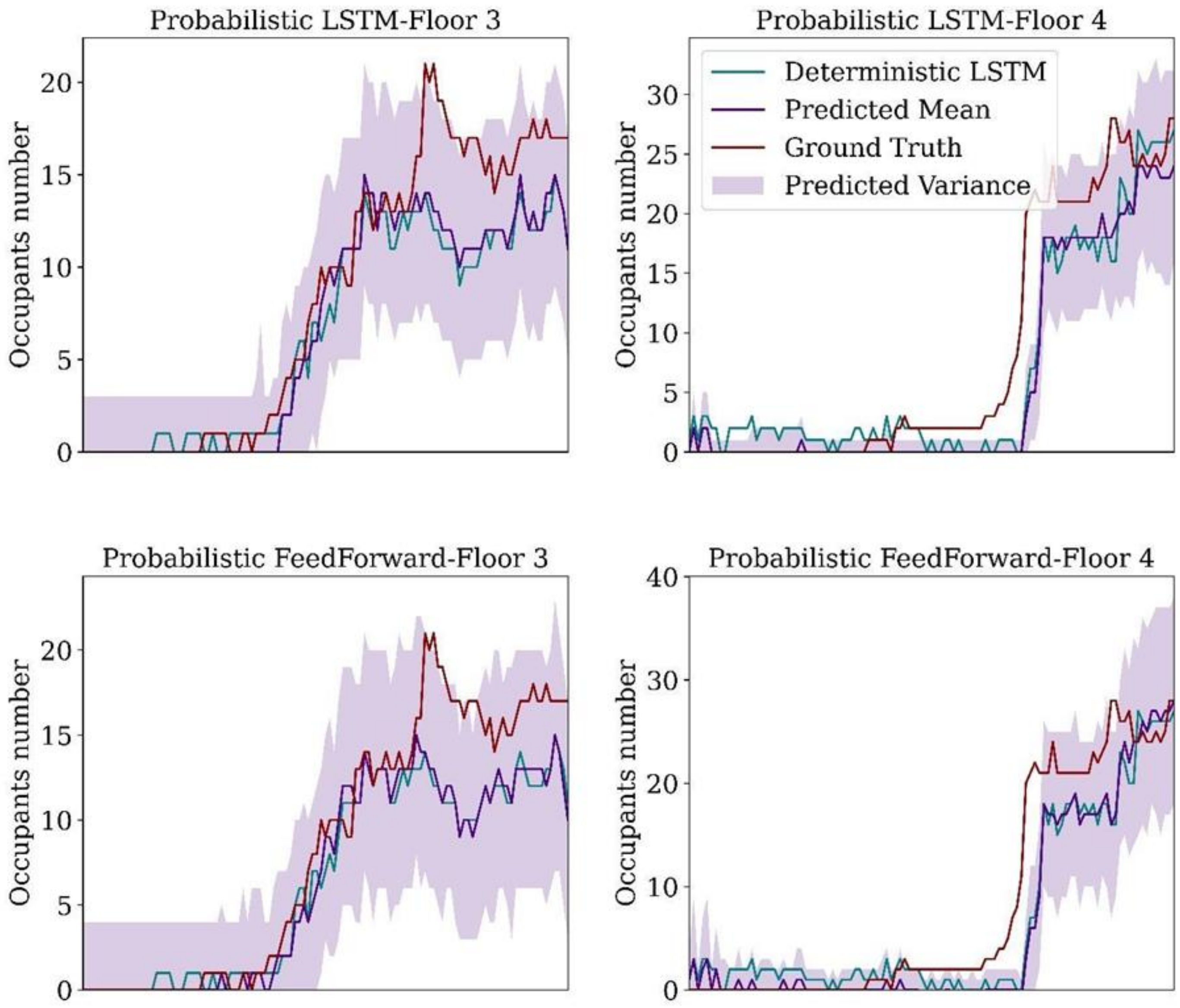
Schematics of models

Developed models

Accuracy of proposed models

Model	Mean Squared Error	Mean Absolute Error	Negative Log-Likelihood
Floor 3			
Probabilistic LSTM	**18**	**2.53**	**4.84**
Probabilistic Feed-Forward	18.95	2.67	2.43
Deterministic LSTM	18.03	2.86	-
Floor 4			
Probabilistic LSTM	**8.2**	**1.9**	**3.4**
Probabilistic Feed-Forward	9.72	2	3.29
Deterministic LSTM	10.24	2.17	-

Predictions by proposed models



EPFL

TEBEL
Thermal Engineering for the Built Environment Laboratory

Thank you

Amirreza Heidari
amirreza.heidari@epfl.ch

Dolaana Kholayg
dolaana.khovalyg@epfl.ch

Verena marie barthelmes
verena.barthelmes@epfl.ch

References:

1. Y. Gal and Z. Ghahramani, “Dropout as a Bayesian approximation: Representing model uncertainty in deep learning,” in 33rd International Conference on Machine Learning, ICML 2016, 2016.

2. B. W. Hobson, D. Lowcay, H. B. Gunay, A. Ashouri, and G. R. Newsham, “Opportunistic occupancy-count estimation using sensor fusion: A case study,” Build. Environ., 2019, doi: 10.1016/j.buildenv.2019.05.032.