Variational Information Bottleneck Model for Accurate Indoor Position Recognition

Weizhu QIAN, Franck GECHTER

CIAD, Université de Technologie de Belfort Montbéliard, UBFC

November 2020

Introduction

Problem Statement

In our task, the input are received signal strength values and the target are the corresponding user coordinates. This task can be regarded as a high dimensional regression task.

Key Idea

We assume that there exists a latent variable Z governing both the input X (WiFi fingerprints) and the target Y (user coordinates). The information Markov chain is: $X \to Z \to Y$.

Assumptions

- **Assumption 1:** There exists a statistic of the input X which is solely sufficient enough to learn the posterior of Z, i.e., $\mathbb{P}(Z|X, Y) = \mathbb{P}(Z|X)$;
- **Assumption 2:** The learned representation Z is solely sufficient enough to learn the likelihood of Y, i.e., $\mathbb{P}(Y|X, Z) = \mathbb{P}(Y|Z)$.

Information Bottleneck Method

The original optimization problem of IB is

$$\max \; I(Z;Y) \tag{1}$$
$$\text{s.t.} \; I(Z;X) \leq I_C \tag{2}$$

where I denotes the mutual information, I_C is the information constrain.

We apply the KKT condition to the above optimization problem, then its corresponding Lagrangian yields

$$I(Z;Y) - \beta\big(I(Z;X) - I_C\big) \tag{3}$$

where β is a positive Lagrangian multiplier governing the optimization trade-off.

IB can be implemented via a supervised deep latent model, i.e., **Variational Information Bottleneck (VIB)**, which can be used to extract the task-related information from the original input so as to reduce the noise.

Variational Information Bottleneck

Since direct computing $I(Z;X)$ and $I(Z;Y)$ via neural networks are tricky, we need to derive variational bounds to approximate them.

First, we derive the variational upper bound for $I(Z;X)$:

$$
\begin{aligned}
I(Z;X) &= \iint p(z,x) \log \left\{ \frac{p(z|x)q(z)}{p(z)q(z)} \right\} dxdz \\
&= \iint p(z,x) \left[\log \left\{ \frac{p(z|x)}{q(z)} \right\} - \log \left\{ \frac{p(z)}{q(z)} \right\} \right] dxdz \\
&= \mathbb{E}_x \left[D_{KL}\big(p(z|x)||q(z)\big) \right] - D_{KL}\Big(p(z)||q(z)\Big) \\
&\leq \mathbb{E}_x \left[D_{KL}\big(p(z|x)||q(z)\big) \right]
\end{aligned} \tag{4}
$$

where $q(z)$ is an uninformative prior, here we can use a standard Gaussian distribution (same as in VAEs).

This is the **upstream task** used to compress the input in the VIB model.

Variational Information Bottleneck

The variational lower bound for $I(Z;Y)$ is derived as follows:

$$
\begin{aligned}
I(Z;Y) &= \mathbb{E}_{p(z,y)}\left[\log\left\{\frac{p(z,y)}{p(z)p(y)}\right\}\right] \\
&= \mathbb{E}_{p(z,y)}\left[\log\left\{\frac{p(y|z)p(z)}{p(z)}\right\} - \log\left\{p(y)\right\}\right] \\
&\geq \mathbb{E}_{p(z,y)}\big[\log\left\{p(y|z)\right\}\big]
\end{aligned}
\tag{5}
$$

We use the mean squared errors to represent $p(y|z)$ (regression task).

This is the **downstream task** used to predict the output of the VIB model.

Finally, combining these two variational bounds, we obtain the loss function of the VIB model solved by the re-parameterization trick and Monte Carlo sampling:

$$
\mathcal{L}(D,\theta,\phi) = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}_{\epsilon_z\sim p(\epsilon_z)}[p_\theta(y_n|f_\phi(x_n,\epsilon_z))] - \beta D_{KL}[p_\phi(z|x_n)||q(z)] \tag{6}
$$

where N is the mini batch size.

VIB Model Structure

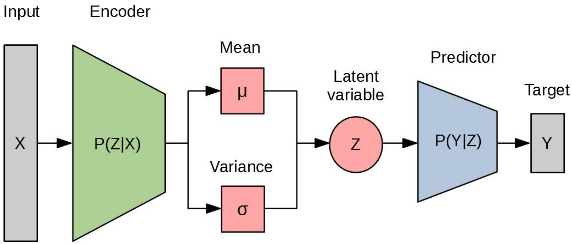
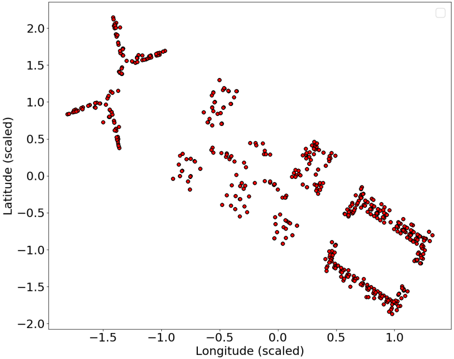


Figure 1: Overview of the VIB-based model.

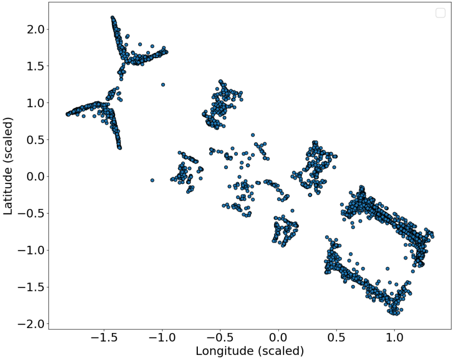
The proposed model consists of an encoder network $p(z|x)$, and a predictor network $p(y|z)$.

Experiment: Prediction Results

We use the UJIIndoor dataset for the validation.

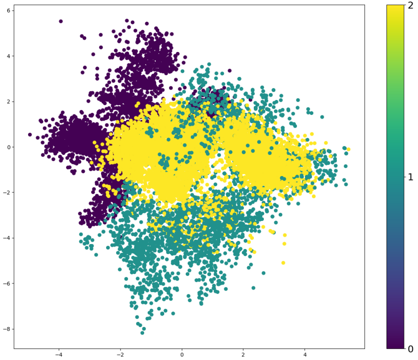


(a) Ground truth.

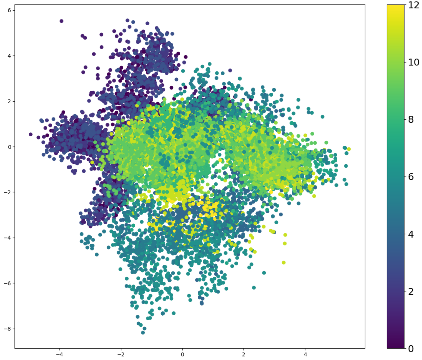


(b) Testing result.

Experiment: Latent Distribution



(a) Latent variables labeled with the building IDs.



(b) Latent variables labeled with the floor IDs.

As it is demonstrated here, the latent distribution is related to the building IDs and floor IDs in some way, which verifies our assumptions.

Experiment: Encoding-Predicting Trade-off

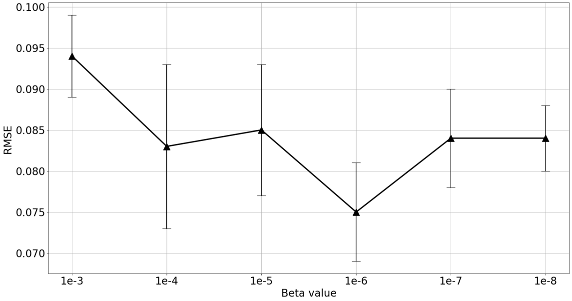


Figure 4: Results with different β values.

β controls the trade-off between the upstream task and the downstream task, thus its value is essential to the performance of the VIB model.

VIB: Comparisons with other Methods

Table 1: Comparison results.

Method	k-NN	GP	DT	RF
RMSE	$0.092 \pm$ 2e-3	$0.252 \pm$ 3e-3	$0.112 \pm$ 3e-3	$0.087 \pm$ 3e-3

Method	MDN-2	MDN-5	BNN	Semi-VAE	Proposed
RMSE	$0.099 \pm$ 3e-4	$0.103 \pm$ 3e-3	$1.033 \pm$ 4e-3	$0.077 \pm$ 4e-3	$0.075 \pm$ 6e-3

Our model outperforms the conventional machine learning models (k-NN, Gaussian Process, Decision Trees and Random Forests) and the deep learning models (Mixture Density Networks, Bayesian Neural Network and the semi-VAE model).

Conclusions

In this work, in order to calculate accurate indoor user location with corresponding WiFi fingerprints, we devised the Variational Information Bottleneck-based model for accurate indoor user location recognition. Our model can extract the useful information from the input so as to improve the accuracy and the final results verify its effectiveness.