

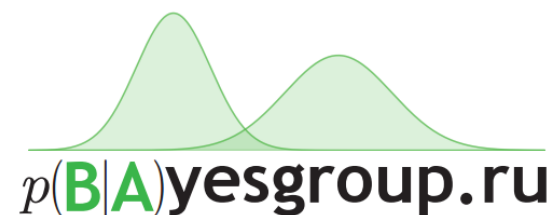
# Глубинное обучение (лекции 7, 8)

Нейробайесовский подход

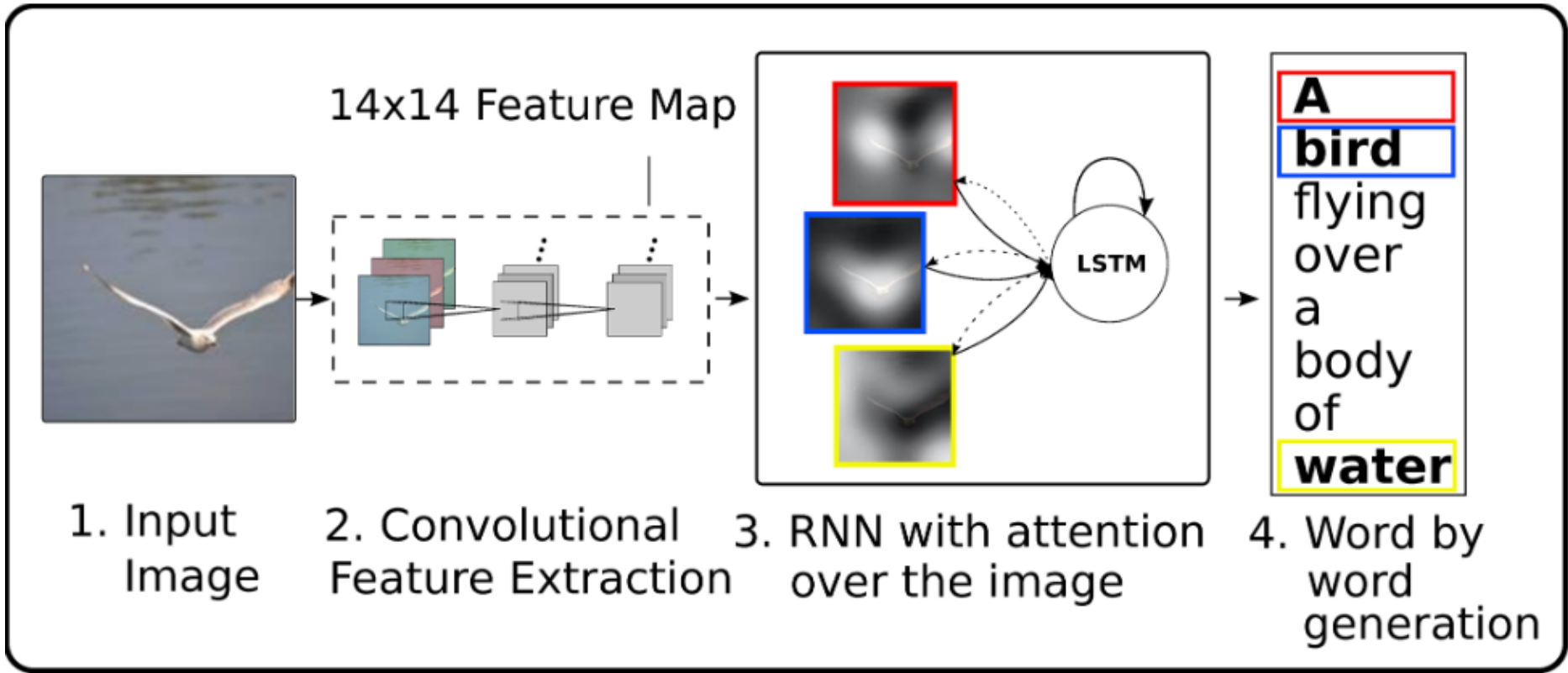
*Д.П. Ветров*

*Руководитель исследовательской группы Байесовских методов*

*Доцент НИУ ВШЭ, факультет компьютерных наук*



# Attention model



K.Xu, J.Ba, R.Kiros, K.Cho, A.Courville, R.Salakhutdinov, R.Zemel, Y.Bengio. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. ICML 2015.

# Soft vs. Hard attention

Soft attention



bird

flying

over

a

body

of

water

Hard attention

# Interpretability

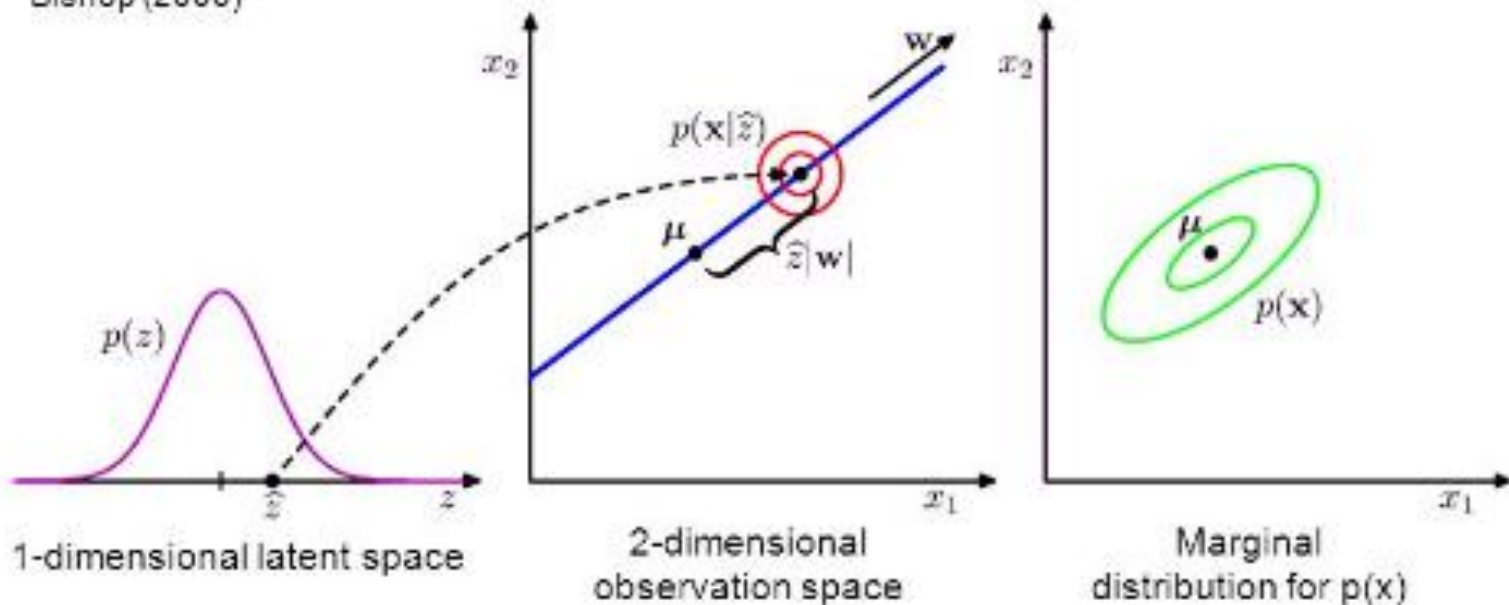


A man is talking on his cell phone while another man watches.

# PCA generative model

## Example of the generative process for PCA

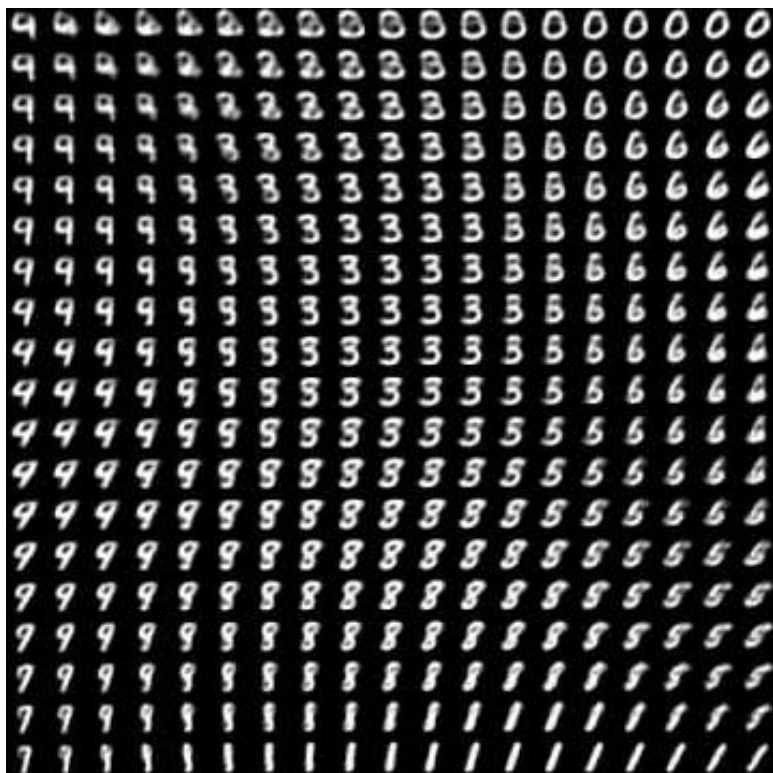
Bishop (2006)



Z = latent variable

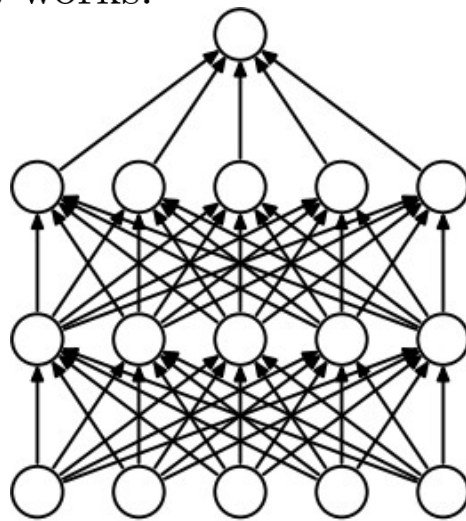
X = observed variable

# Variational auto-encoder

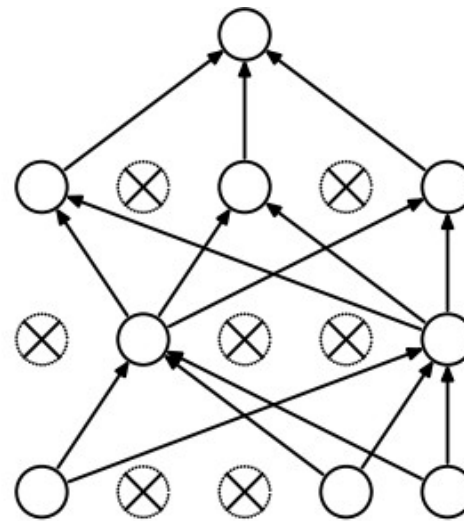


# Dropout

- Proposed by Geoffrey Hinton's group (Srivastava14)
- Nullifies the outputs of randomly selected neurons at each iteration of training
- Purely heuristic procedure for preventing overfitting
- But it works!



(a) Standard Neural Net



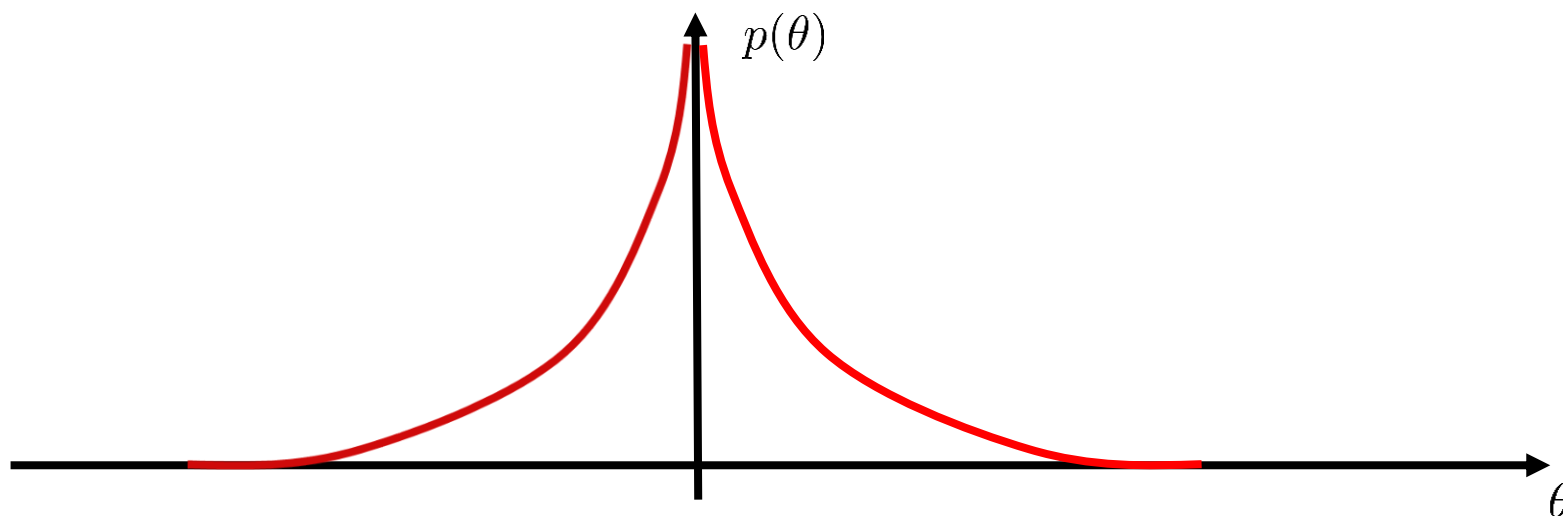
(b) After applying dropout.

# Variational Dropout

- In December 2015 new techniques of variational dropout was suggested (Kingma15)
- It was shown that dropout corresponds to Bayesian inference with special improper prior over the weights  $\theta$

$$p(\theta) \propto \frac{1}{|\theta|}$$

- This is so-called **scale-invariant prior** which penalizes the precision of  $\theta$





# Artistic Style



# Deep RL

