Closing Remarks

Wilker Aziz

Universiteit van Amsterdam
w.aziz@uva.nl

May 24, 2017

NLP2

This course covered

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation
- A bit of math: constrained optimisation, exponential families, semirings

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation
- A bit of math: constrained optimisation, exponential families, semirings

Our latent variables:

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation
- A bit of math: constrained optimisation, exponential families, semirings

Our latent variables:

- alignments (e.g. IBM1)

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation
- A bit of math: constrained optimisation, exponential families, semirings

Our latent variables:

- alignments (e.g. IBM1)
- parameters (e.g. Bayesian IBM1)

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation
- A bit of math: constrained optimisation, exponential families, semirings

Our latent variables:

- alignments (e.g. IBM1)
- parameters (e.g. Bayesian IBM1)
- trees (e.g. LV-CRF for ITGs)

NLP2

This course covered

- Modelling techniques that power structure prediction with latent variables in NLP: mixture models, CRFs, VAE
- Major learning techniques/algorithms: MLE and (approximate) posterior inference by VI
- We also covered some background on logical problems
  - parsing with CFGs and generalisations
  - hypergraph algorithms: inside, outside, expectation
- A bit of math: constrained optimisation, exponential families, semirings

Our latent variables:

- alignments (e.g. IBM1)
- parameters (e.g. Bayesian IBM1)
- trees (e.g. LV-CRF for ITGs)
- continuous representations (e.g. VAE)

What else is there?

Stuff you should learn about

- Other deep generative models: GANs
- Bayesian NNs
- Bayesian nonparametric models
- Gaussian Processes
- Global optimisation
- Sampling: MC and MCMC

Beyond

I'll be offering a course on Bayesian NNs (stay tuned!)

- hottest topics in deep generative modelling
- e.g. Bayesian GANs and Bayesian VAEs
- a dry run might happen as a June course
  (email me if you are interested)
- lots of theory, applications somewhat toy-ish
- small group of students with good ML/statistics background

Beyond

I'll be offering a course on Bayesian NNs (stay tuned!)

- hottest topics in deep generative modelling
- e.g. Bayesian GANs and Bayesian VAEs
- a dry run might happen as a June course (email me if you are interested)
- lots of theory, applications somewhat toy-ish
- small group of students with good ML/statistics background

I'm offering practical projects

- e.g. 6EC courses (1-3 students), dissertation
- more applied problems
- findings typically lead to a publication

Beyond

I'll be offering a course on Bayesian NNs (stay tuned!)

- hottest topics in deep generative modelling
- e.g. Bayesian GANs and Bayesian VAEs
- a dry run might happen as a June course
  (email me if you are interested)
- lots of theory, applications somewhat toy-ish
- small group of students with good ML/statistics background

I'm offering practical projects

- e.g. 6EC courses (1-3 students), dissertation
- more applied problems
- findings typically lead to a publication

Important

- None of it is plain supervised learning
- none of it is 100% frequentist
- most of it about deep generative models

Projects

Examples:

1. Bayesian attention: model attention weights as a random variable.
2. Joint modelling for NMT: model a joint distribution and use monolingual data as semi-supervision.
3. Mixture model for NMT: a mixture model to alleviate the LM bias of NMT.
4. VAE+CRF project: discrete latent space VAE with approximation given by a CRF.

What's next?

First,

1. you finish project 3

What's next?

First,

1. you finish project 3
2. we grade project 3

What's next?

First,

1. you finish project 3
2. we grade project 3

Then,

- I fly to a warm place and enjoy a bit of summer

What's next?

First,

1. you finish project 3
2. we grade project 3

Then,

- I fly to a warm place and enjoy a bit of summer
- and hopefully you too :D

What's next?

First,

1. you finish project 3
2. we grade project 3

Then,

- I fly to a warm place and enjoy a bit of summer
- and hopefully you too :D

Finally,

- feel free to contact me in the future