Generalized Zero Shot Learning for Intent Classification and Slot Filling

Korolev Alexey
scientific advisors: Bondarenko Ivan, Batura Tatyana

Big Data & AI
Novosibirsk State University

22 December

Content

Introduction to DST

Nowadays many services implement in their works dialogue system because it helps to provide services at any time of the day without additional labour costs. DST is a Core component in today's task-oriented dialogue systems, maintains a user's intentional states through the course of a dialogue.

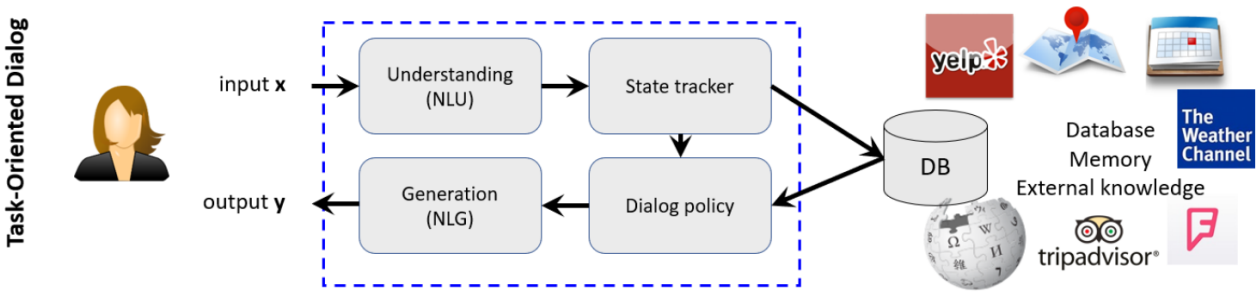

Figure: Dialogue system

Introduction to DST

During the dialogue, we recognize the user intends and fill the appropriate slot.
For different services, we have different intends and slots.
So if we include new service we need to provide train dialogues and train again system.

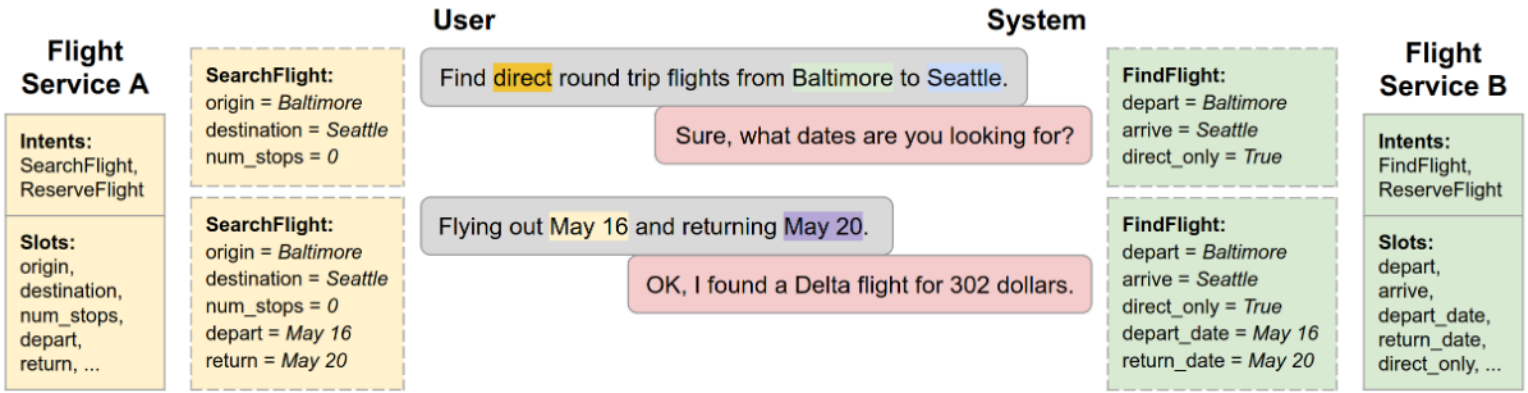

Figure: Example of shema dialogue system.

GZSL

Generalized Zero-shot learning aims to recognise objects whose class may not have been seen during training.
For classes, we have some description. So GZSL methods trying to extract knowledge between class description and class example.

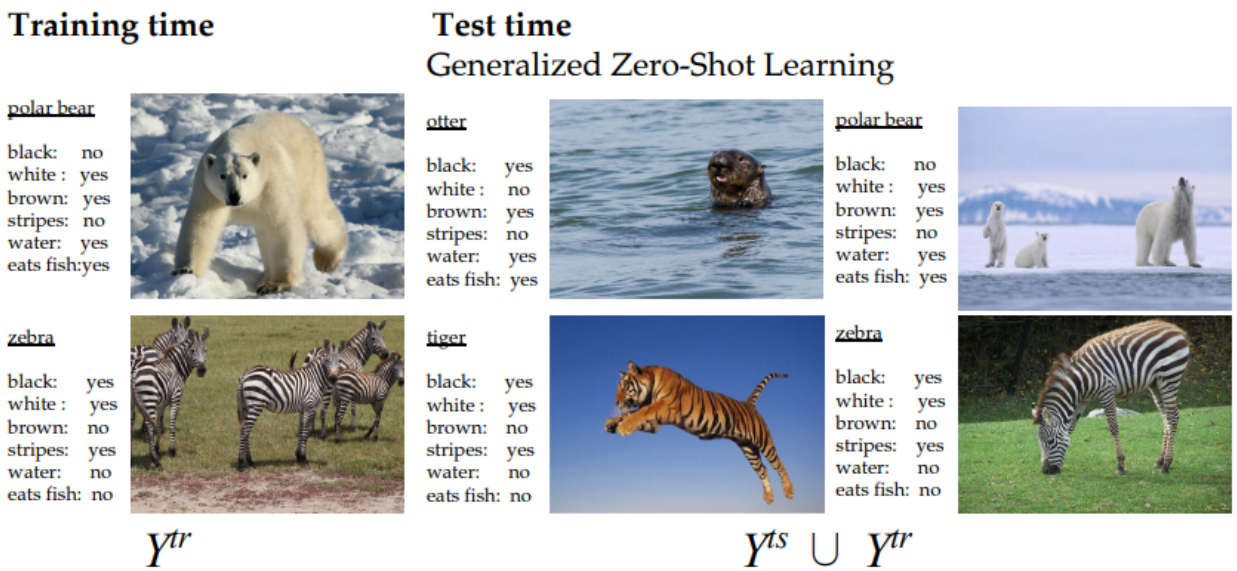

Figure: Flow visualisation[1]

[1] B. Schiele Y. Xian C. H. Lampert and Z. Akata. "Zero-Shot Learning - A Comprehensive Evaluation of the Good, the Bad and the Ugly". In: (2018). URL:

ZeroShotEval

Our command creating a tool for Zero-shot task. Using this system you can perform zls task or evaluate your architecture on wild used datasets.

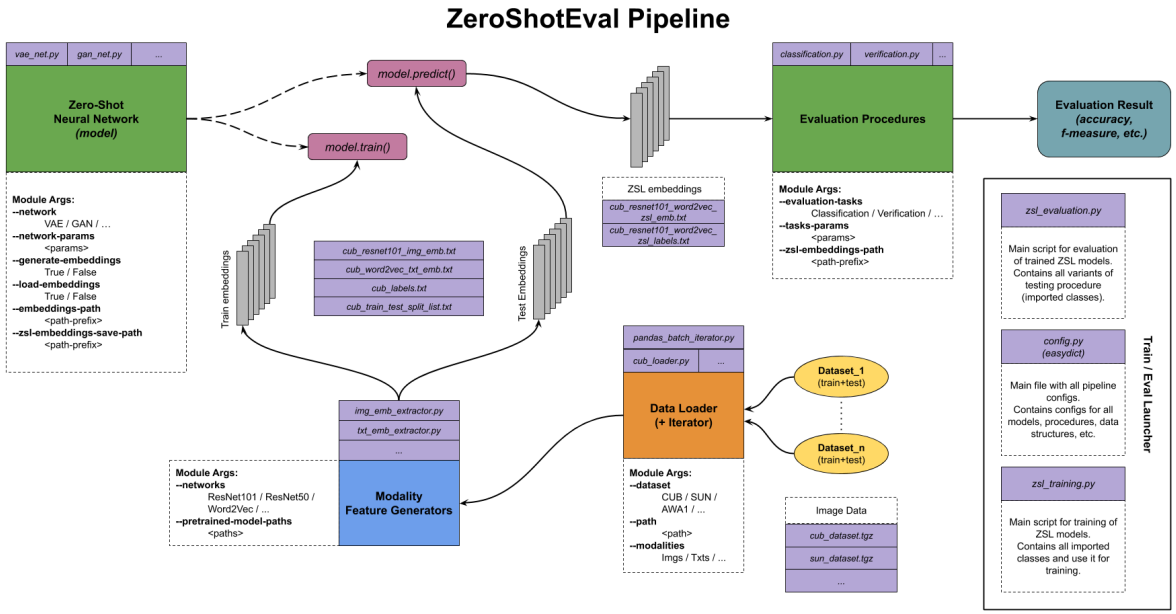

Figure: Strurcure of Zero-shot eval system[2]

[2] **https://github.com/ZSLresearch-team/ZeroShotEval**.

GZSL Nets

There's a lot of different approaches to perform GZSL task.
In ZeroShortEval inplement 2 nets: CADA-VAE and CLSWGAN.
CADA-VAE: based on Variation auto encoder.[3]
LISGAN: based on GAN[4]
for input to this nets, we take picture embedding(resnet-101) and class embedding. output: ZSL-embedding.
Also, there are Bayesian networks, knowledge graph etc.

[3] S. Sinha T. Darrell Z. Akata E. Schonfeld S. Ebrahimi. "Generalized Zero- and Few-Shot Learning via Aligned Variational Autoencoders". In: (2019). URL: https://arxiv.org/pdf/1812.01784.pdf.

[4] B. Schiele Z. Akata Y. Xian T. Lorenz. "Feature Generating Networks for Zero-Shot Learning". In: (2018). URL: https://arxiv.org/pdf/1712.00981v2.pdf.

BERT approaches

BERT is a popular architecture for many NLP tasks. There are o lot of works in DST with BERT. For example in this work with architecture similar to this work..[5]

$[h_0, \ldots, h_T] = BERT([e_1, \ldots, e_T]); y_i = softmax(Wh_i + b);$

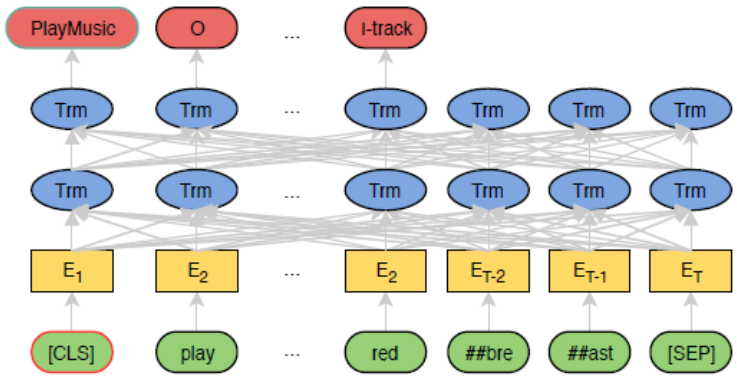

Figure: A high-level view of the observed model. The input query is "play the song little robin redbreast".

[5] W. Wang Q. Chen Z. Zhuo. "BERT for Joint Intent Classification and Slot Filling". In: (2019). URL: https://arxiv.org/pdf/1902.10909.pdf.

Datasets

During literature reviews next datasets is frequent occurrence: DSTC 2, SNIPS.
Also occure: XSchema, MultiWOZ 2.1., ect
Also some special Dataset DSTC-8 (SGP-DST).[6]Special new dataset for ZSL approach in DST. It contain different services with different intents.

[6] S. Sunkara R. Gupta P. Khaitan A. Rastogi X. Zang. "Schema-Guided Dialogue State Tracking Task at DSTC8". In: (2020). URL: https://arxiv.org/pdf/2002.01359.pdf.

Metrics

1. Active Intent Accuracy: The fraction of user turns for which the active intent has been correctly predicted.
2. Harmonic mean between seen and unseen classes accuracy.[7]

So for intent classification we use accuracy.

[7] Y. Xian and Akata., "Zero-Shot Learning - A Comprehensive Evaluation of the Good, the Bad and the Ugly".

Results

We choose DSTC8 dataset.

1. We parse all dialogues and choose utterance, in which appeared intent.
2. Encode this intent using RoBERTA sentence encoder.
3. Encode intent description RoBERTA sentence encoder
4. Train LisGAN and CADAVAE nets on standard parameters.

Our created dataset contains 40 seen classes, 6 unseen.
Train 49892 sample test 11701 sample (10469 seen, 1232 unseen)
Size of utterance embedding 768
Our results for now:
CADAVAE: Seen 73.87 Unseen: 50.77 H: 60.18 Intent Accuracy: 72.43
LISGAN: Seen: 86.08, Unseen: 4.87, FC H: 9.22 Intent Accuracy: 77.5

What next

1. Continue develop ZeroShotEval.
2. Try fine-tune parameters.
3. Try another encoder (XMLRoBETA)
4. Add service description to intent description.

References

- A. Rastogi X. Zang, S. Sunkara R. Gupta P. Khaitan. "Schema-Guided Dialogue State Tracking Task at DSTC8". In: (2020). URL: https://arxiv.org/pdf/2002.01359.pdf.
- E. Schonfeld S. Ebrahimi, S. Sinha T. Darrell Z. Akata. "Generalized Zero- and Few-Shot Learning via Aligned Variational Autoencoders". In: (2019). URL: https://arxiv.org/pdf/1812.01784.pdf.
- Q. Chen Z. Zhuo, W. Wang. "BERT for Joint Intent Classification and Slot Filling". In: (2019). URL: https://arxiv.org/pdf/1902.10909.pdf.
- Y. Xian C. H. Lampert, B. Schiele and Z. Akata. "Zero-Shot Learning - A Comprehensive Evaluation of the Good, the Bad and the Ugly". In: (2018). URL: https://arxiv.org/pdf/1707.00600.pdf.
- Y. Xian T. Lorenz, B. Schiele Z. Akata. "Feature Generating Networks for Zero-Shot Learning". In: (2018). URL: https://arxiv.org/pdf/1712.00981v2.pdf.

Thank you for attenrion.