

Vysoká škola ekonomická v Praze
Fakulta informatiky a statistiky



Variační autoenkodér a úlohy pozorování v latentním prostoru

BAKALÁŘSKÁ PRÁCE

Studijní program: Aplikovaná informatika

Autor: Tomáš Faltejsek

Vedoucí práce: Ing. Ondřej Vadinský, Ph.D.

Konzultant práce: full consultant's name (incl. degrees)

Praha, květen 2023

Poděkování

Thanks.

Abstract

Jedním z předních rysů lidské inteligence je intuice a schopnost představovat si nové objekty. Variační autoenkodér je inovací na poli pravděpodobnostních modelů, umožňující architekturu modelů schopných syntézy zcela nových dat s využitím pozorování atributů v latentním prostoru. Teoretická charakteristika a možnosti využití variačního autoenkodéru jsou předmětem této bakalářské práce.

Keywords

keyword, important term, another topic, and another one

Obsah

Úvod	9
1 Východiska variačního autoenkodéru	10
1.1 Neuronové sítě	11
1.1.1 Universal approximation theorem	11
1.2 Redukce dimenzionality	12
1.2.1 The Curse of Dimensionality	12
1.2.2 Analýza hlavních komponent	12
1.3 Autoenkodér	13
1.3.1 Historický pohled	13
1.3.2 Od Analýzy hlavních komponent po Autoenkodér	13
1.4 Pravděpodobnostní modely	14
1.5 Kullback–Lieblerova divergence	15
1.6 Modely využívající latentních proměnných // Latent Variable Models	16
2 Variační autoenkodér	17
2.1 ELBO	17
2.2 Reparametrizační trik	17
2.3 Formalizace	17
2.4 Model neuronové sítě	17
2.5 Nedostatky a omezení	17
2.6 Rozšíření a aktuální stav poznání	17
2.7 Pozorování v latentním prostoru	17
3 Úlohy pozorování v latentním prostoru	18
3.1 Generativní modelování obrazových dat	18
3.2 Rekonstrukce obrazových dat	18
3.3 Interpolace vět	18
3.4 Detekce anomálií	18
3.5 Syntéza tabulárních dat	18
3.6 Komprese	18
4 Experimenty s modelem variačního autoenkodéru	19
4.1 Generativní modelování obrazových dat	19
4.1.1 Vymezení problémové oblasti	19
4.1.2 Datová sada a předzpracování	19
4.1.3 Nastavení experimentu	19
4.1.4 Návrh modelu	19
4.1.5 Evaluace	19
4.1.6 Diskuze	19

4.2 Interpolace vět	19
Závěr	20
A Zdrojové kódy modelů	22

Seznam obrázků

Note: Add a list of figures if the number of figures in the thesis text exceeds 20. A list of diagrams is applicable only if the author distinguishes between a figure and a diagram. The list of diagrams is included if the number of diagrams exceeds 20. This thesis template does not distinguish between a figure and a diagram.

Seznam tabulek

Note: Add a list of tables if the number of tables used in the thesis exceeds 20.

Seznam použitých zkratek

BCC Blind Carbon Copy

CC Carbon Copy

CERT Computer Emergency Response
Team

CSS Cascading Styleheets

DOI Digital Object Identifier

HTML Hypertext Markup Language

REST Representational State Transfer

SOAP Simple Object Access Protocol

URI Uniform Resource Identifier

URL Uniform Resource Locator

XML eXtended Markup Language

Note: Add a list of abbreviations if the number of abbreviations used in the thesis exceeds 20 and the abbreviations used are not common.

Úvod

Introduction is a compulsory part of the bachelor's / diploma thesis. The introduction is an introduction to the topic. It elaborates the chosen topic, briefly puts it into context (there may also be a description of the motivation to write the work) and answers the question why the topic was chosen. It puts the topic into context and justifies its necessity and the topicality of the solution. It contains an explicit goal of the work. The text of the thesis goal is identical with the text that is given in the bachelor's thesis assignment, ie with the text that is given in the InSIS system and which is also given in the Abstract section.

Part of the introduction is also a brief introduction to the process of processing the work (a separate part of the actual text of the work is devoted to the method of processing). The introduction may also include a description of the motivation to write the work.

The introduction to the diploma thesis must be more elaborate - this is stated in more detail in the Requirements of the diploma thesis within the Intranet for FIS students.

Here are some sample chapters that recommend how a bachelor's / master's thesis should be set. They primarily describe the use of the L^AT_EX template, but general advice will also serve users of other systems well.

1. Východiska variačního autoenkodéru

1.1 Neuronové sítě

1.1.1 Universal approximation theorem

1.2 Redukce dimenzionality

1.2.1 The Curse of Dimensionality

1.2.2 Analýza hlavních komponent

1.3 Autoenkodér

1.3.1 Historický pohled

1.3.2 Od Analýzy hlavních komponent po Autoenkodér

1.4 Pravděpodobnostní modely

1.5 Kullback–Lieblerova divergence

1.6 Modely využívající latentních proměnných // Latent Variable Models

2. Variační autoenkodér

2.1 ELBO

2.2 Reparametrizační trik

2.3 Formalizace

2.4 Model neuronové sítě

2.5 Nedostatky a omezení

2.6 Rozšíření a aktuální stav poznání

2.7 Pozorování v latentním prostoru

3. Úlohy pozorování v latentním prostoru

3.1 Generativní modelování obrazových dat

3.2 Rekonstrukce obrazových dat

3.3 Interpolace vět

3.4 Detekce anomálií

3.5 Syntéza tabulárních dat

3.6 Komprese

4. Experimenty s modelem variačního autoenkodéru

4.1 Generativní modelování obrazových dat

4.1.1 Vymezení problémové oblasti

4.1.2 Datová sada a předzpracování

4.1.3 Nastavení experimentu

4.1.4 Návrh modelu

4.1.5 Evaluace

4.1.6 Diskuze

4.2 Interpolace vět

Závěr

The conclusion is a mandatory part of the bachelor's / diploma thesis. It contains a summary of the work and comments on the degree of fulfillment of the goal, which was set in the work, or summarizes the answers to the questions that were asked in the introduction.

The conclusion to the diploma thesis must be more elaborate - this is stated in more detail in the Requirements of the diploma thesis within the Intranet for FIS students.

The conclusion is perceived as a chapter, which begins on a separate page and is called the conclusion. The name Conclusion is not numbered. The text of the conclusion itself is divided into paragraphs.

Přílohy

A. Zdrojové kódy modelů