

KAZAKH NATIONAL UNIVERSITY NAMED AFTER AL-FARABI

D.R. Rakhimova, Zh.M. Zhumanov

MACHINE TRANSLATION TECHNOLOGIES

Electronic Textbook

Almaty
"Kazakh University"
2022

UDC 80/81
LBC 81.2
R15

*Recommended for publication by the Academic Council
Faculty of Information Technology and RISO
of Al-Farabi KazNU (Protocol No. 1 of 9.10.2021)*

Reviewers:

O.Zh. Mamyrbayev – PhD., associate professor. Deputy General
Director RSE "Institute of Information and computing technologies"
M.E. Mansurova – PhD., assistant professor,
Head of the department "Artificial Intelligence and Big Data"
Al-Farabi KazNU

Rakhimova D.R., Zhumanov Zh.M.

R 15 Machine translation technologies: electronic textbook /
D.R. Rakhimova, Zh.M. Zhumanov. – Almaty: Kazakh
University, 2022. – 180 p.
ISBN 978-601-04-6046-1

The tutorial describes systems and modern methods of machine translation for natural language. The manual can be used for undergraduate, graduate and doctoral students of the specialties "Information Systems", "Information Security Systems", "Informatics". This tutorial can be used for graduate and master's theses in computational linguistics and artificial intelligence.

Acknowledgments.

This work was performed by the grant Project IRN AP 08052421 of Ministry of Science and Higher Education of the Republic of Kazakhstan and Erasmus CLASS Project "Development of the interdisciplinary master program on Computational Linguistics at Central Asian universities" Number 585845-EPP-1-2017-1-ES-EPPKA2-CBHE-JP

**UDC 80/81
LBC 81.2**

Contents

| | |
|--|-----|
| Introduction | 4 |
| Chapter 1. Overview and Architectures of machine translation systems | 5 |
| 1.1. Overview of machine translation problems | 5 |
| 1.2. Architectures of machine translation systems | 11 |
| 1.3 Practical work | 17 |
| Chapter 2. Rule-based machine translation | 19 |
| 2.1. Basic concepts of RBMT. Overview of resources needed for RBMT | 19 |
| 2.2. RBMT models. Apertium machine translation platform | 28 |
| 2.3. Practical work | 41 |
| Chapter 3. Statistical machine translation | 43 |
| 3.1. Basic concepts of SMT. Overview of resources needed for SMT | 43 |
| 3.2. SMT models: Word-Based Models. Phrase-Based Models. Decoding | 49 |
| 3.3. Practical work | 72 |
| Chapter 4. SMT models | 73 |
| 4.1. Language Models. Tree-Based Models | 73 |
| 4.2. Moses statistical machine translation system | 95 |
| 4.3. Practical work | 110 |
| Chapter 5. Hybrid machine translation | 112 |
| 5.1. Applications of Machine Translation | 112 |
| 5.2. Practical work | 119 |
| Chapter 6. Neural machine translation | 120 |
| 6.1. The basic concepts, principle of operation and review of approaches | 120 |
| 6.2. Neural Machine Translation with TensorFlow | 130 |
| 6.3. Practical work | 143 |
| Chapter 7. Neural machine translation models | 147 |
| 7.1. Neural Machine Translation using an Attention Mechanism | 147 |
| 7.2. Neural machine translation using a Transformer model | 155 |
| 7.3. Practical work | 162 |
| Chapter 8. Evaluation of Machine Translation Systems | 166 |
| 8.1. Review of Evaluation methods of Machine Translation Systems | 166 |
| 8.2. Practical work | 169 |
| Chapter 9. Post-editing system in machine translation | 173 |
| 9.1 Post-Editing System Overview and Modules | 173 |
| 9.2 Test runs and evaluation of the translation quality of the post-editing system prototype | 178 |