Didar Kanalbay
Group: sis-2224

PRACTICAL WORK No. 7-8
Topic: Designing the Structure of a Thesis and the Logic of Scientific Research

Research topic:
Creating a multiclass model for detecting destructive web content in social
networks and instant messengers using machine learning.

## 1. Purpose of the Practical Work

The purpose of this practical work is to design a logically consistent thesis
structure where the scientific problem, research goal, objectives, methods, and
expected results are explicitly connected. The developed structure must support
step-by-step scientific reasoning from problem analysis to model development,
experimental validation, and practical recommendations.

## 2. Requirements for the Thesis Structure

The thesis structure is designed to:
- reflect the logic of scientific research;
- provide a consistent unfolding of the topic from theory to practice;
- correspond to the stated research objectives;
- include theoretical, methodological, and experimental-practical parts;
- avoid duplication and ensure clear differentiation of chapter roles.

## 3. Developed Complete Table of Contents (for this topic)

INTRODUCTION
- Relevance of the topic
- Problem statement and contradiction analysis
- Research goal and objectives
- Object and subject of research
- Scientific novelty and practical significance
- Research methods
- Thesis structure overview

CHAPTER 1. THEORETICAL FOUNDATIONS OF DETECTING DESTRUCTIVE WEB CONTENT
1.1 Analysis of the subject area
    1.1.1 Definition and taxonomy of destructive web content
    1.1.2 Specific features of social networks and instant messengers
    1.1.3 Legal, ethical, and platform constraints of content moderation
1.2 Review of methods and approaches
    1.2.1 Classical machine learning for text classification
    1.2.2 Deep learning and transformer-based NLP models
    1.2.3 Multilingual and domain-adapted approaches
1.3 Comparative analysis of existing solutions
    1.3.1 Academic datasets and benchmark systems
    1.3.2 Industrial moderation solutions and available APIs
    1.3.3 Strengths, weaknesses, and unresolved gaps
1.4 Conclusions for Chapter 1

CHAPTER 2. METHODOLOGY AND DEVELOPMENT OF THE MULTICLASS SOLUTION
2.1 Problem statement

4. Connection Between Chapters and Research Objectives

Research Goal:

To develop and experimentally validate a multiclass machine learning model for detecting destructive web content in social networks and instant messengers,

with improved classification quality and practical applicability.

Research Objectives:
O1. Analyze the domain and define destructive content classes.
O2. Review and compare existing machine learning approaches.
O3. Formulate the multiclass problem and select methods and tools.
O4. Develop the model, training algorithm, and system architecture.
O5. Conduct experiments and evaluate results against baselines/analogues.
O6. Generalize findings and formulate conclusions and recommendations.

Objective-to-Structure Mapping:
- O1 -> Chapter 1 (1.1, 1.3): domain analysis and taxonomy definition.
- O2 -> Chapter 1 (1.2, 1.3): method review and comparative gap analysis.
- O3 -> Chapter 2 (2.1, 2.2): formalization, metrics, and methodology choice.
- O4 -> Chapter 2 (2.3, 2.4): model design, algorithm, and architecture.
- O5 -> Chapter 3 (3.1-3.5): experiment design, results, and benchmarking.
- O6 -> Conclusion + chapter conclusions: synthesis and recommendations.

5. Table Linking Objectives and Structure

| Research objective | Thesis section(s) | Methods | Expected results |
|--------------------|-------------------|---------|------------------|
| O1: Domain analysis and class taxonomy | 1.1, 1.3 | Literature review, domain analysis, content typology | Formal taxonomy of destructive content and problem boundaries |
| O2: Review of ML approaches | 1.2, 1.3 | Systematic review, comparative analysis | Justifi selection criteria for candidate methods |
| O3: Problem formalization and method selection | 2.1, 2.2 | Task formalization, metric design, tool evaluation | Defined I/O model, metrics, dataset strategy, software stack |
| O4: Model and architecture development | 2.3, 2.4 | Algorithm design, prototyping, engineering design | Implemented multiclass model and deployable architecture |
| O5: Experimental validation | 3.1-3.5 | Controlled experiments, statistical evaluation, benchmark comparison | Quantitative evidence of model quality and limitations |
| O6: Final synthesis | 3.6, Conclusion | Generalization, interpretation, recommendation des | Scientifically grounded conclusions and practical guidance |

6. Content of Each Main Section (What to Write There)

Introduction:
- Explain relevance: growth of harmful/destructive content in online platforms.
- State scientific problem: need accurate multiclass detection, not only binary.
- Define goal, objectives, object, subject, novelty, and practical significance.

Chapter 1 (Theoretical):
- Build terminology and class taxonomy for destructive content.
- Compare classical ML, deep learning, and transformer approaches.
- Identify limitations of current approaches and justify research niche.

Chapter 2 (Methodology and Development):
- Provide exact formal problem statement and assumptions.
- Describe dataset construction and annotation quality assurance.
- Justify chosen model architecture and training strategy.
- Present full solution workflow from data ingestion to inference.

Chapter 3 (Experimental):
- Describe reproducible experiment setup and metrics.

- Show baseline and proposed model results, including per-class details.
- Analyze errors, robustness, and reasons behind improvements/weaknesses.
- Compare with analogues and discuss practical deployment implications.

Conclusion:
- Summarize contributions against each objective.
- Report practical value for moderation systems.
- Identify limitations (e.g., domain drift, multilingual ambiguity) and future work.


## 7. Logical Model of the Thesis Structure

Scientific problem
  -> Goal of research
    -> Objectives (O1..O6)
      -> Chapters and sections (Introduction, Ch1, Ch2, Ch3, Conclusion)
        -> Methods (review, modeling, experiments, statistics)
          -> Results (metrics, comparisons, analyses)
            -> Conclusions and practical recommendations

Expanded logical chain for this topic:
Problem of destructive content moderation in digital communication
  -> Need for multiclass machine learning detection model
    -> Domain analysis + method review + formalization
      -> Model and architecture development
        -> Experimental validation against baselines/analogues
          -> Scientifically justified conclusions and implementation guidance


## 8. Typical Mistakes and How This Structure Avoids Them

1) Lack of connection between objectives and chapters:
  - Solved by explicit O1..O6 mapping to specific sections.
2) Overly general chapter/section titles:
  - Solved by topic-specific titles (taxonomy, imbalance handling, ablation).
3) Inconsistency between theoretical and practical parts:
  - Solved by bridge from Chapter 1 gaps to Chapter 2 model design decisions.
4) Duplication of content:
  - Solved by clear role separation: theory (Ch1), development (Ch2),
    validation (Ch3), synthesis (Conclusion).


## Final Note

The proposed thesis structure is logically coherent and methodologically aligned
with the research topic "Creating a multiclass model for detecting destructive
web content in social networks and instant messengers using machine learning".
It ensures traceability from scientific problem to validated practical results.