UNIVERSITAS
GADJAH MADA
**A Deep Learning Approach with Neural Networks to Recognise Javanese Dialects**
ALVIN JANUAR RAMADAN, Sigit Priyanta, S.Si., M.Kom., Dr.
**Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/**

# Contents

**APPROVAL PAGE** i

**PLAGIARISM STATEMENT** i

**MOTTO** ii

**DEDICATION** iv

**FOREWORDS** v

**TABLE OF CONTENTS** x

**LIST OF TABLES** xi

**LIST OF FIGURES** xii

**ACRONYMS** xiii

**GLOSSARY** xvi

**ABSTRAK** xix

**ABSTRACT** xx

**1 Introduction** 1
1.1 Background . . . 1
1.2 Research Problem . . . 3
1.3 Problem Statement . . . 4
1.4 Research Scope . . . 5
1.5 Research Objectives . . . 5
1.6 Research Advantages . . . 6

**2 Literature Review** 7
2.1 Preface . . . 7
2.2 Data Augmentation . . . 8
2.3 Feature Extraction . . . 8
2.4 Modelling and Classification . . . 11

UNIVERSITAS
GADJAH MADA
**A Deep Learning Approach with Neural Networks to Recognise Javanese Dialects**
ALVIN JANUAR RAMADAN, Sigit Priyanta, S.Si., M.Kom., Dr.
**Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/**

2.5 Neural Network Architectures . . . 12
2.6 Research Comparison . . . 15

**3 Theory 17**
3.1 Corpus . . . 17
3.1.1 Speech corpus . . . 18
3.2 Language / Dialect Identification . . . 18
3.3 Mel Filter Bank . . . 19
3.4 Mel Frequency Cepstral Coefficients . . . 20
3.5 Model . . . 22
3.5.1 Model training . . . 22
3.5.2 Model validation . . . 22
3.6 Neural Network . . . 23
3.6.1 Multi-layer perceptrons . . . 24
3.6.2 Recurrent Neural Network . . . 25
3.6.3 Convolutional Neural Network . . . 25
3.7 Short-time Fourier Transform . . . 28

**4 Research Methodology 31**
4.1 Research Description . . . 31
4.2 Tools and Materials . . . 31
4.2.1 Tools . . . 31
4.2.2 Dataset . . . 32
4.3 Research Phases . . . 36
4.3.1 Literature study . . . 36
4.3.2 Dataset and features collection . . . 37
4.3.3 Model structural design . . . 38
4.3.4 Model training and validation . . . 39
4.3.5 Model comparison . . . 40
4.3.6 Evaluation . . . 40
4.3.7 Final report . . . 44

**5 Implementation 45**
5.1 Dataset . . . 45
5.2 Feature Extraction . . . 46
5.2.1 Wave file loading . . . 46
5.2.2 Frequency spectrogram . . . 47
5.2.3 Mel-scale spectrogram . . . 48
5.2.4 Mel Frequency Cepstral Coefficients . . . 49
5.3 Feature Preprocessing . . . 49
5.3.1 Pre-emphasis . . . 50
5.3.2 Low/High-pass filters . . . 50
5.3.3 Voice activity detection . . . 51
5.3.4 Segment splitting . . . 52
5.3.5 Cepstral mean-variance normalisation . . . 53

UNIVERSITAS
GADJAH MADA
**A Deep Learning Approach with Neural Networks to Recognise Javanese Dialects**
ALVIN JANUAR RAMADAN, Sigit Priyanta, S.Si., M.Kom., Dr.
**Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/**

5.4 Model Conception and Training . . . . . 54
5.4.1 Model architectures . . . . . 54
5.4.2 Model training . . . . . 61

**6 Results and Discussion 64**
6.1 Model Performances . . . . . 64
6.1.1 Mel-scale spectrogram . . . . . 64
6.1.2 Mel frequency cepstral coefficients . . . . . 66
6.1.3 Frequency spectrogram . . . . . 68
6.2 Per-dialect Performances . . . . . 71
6.2.1 Mel-scale spectrogram . . . . . 72
6.2.2 Frequency spectrogram . . . . . 75
6.2.3 Mel frequency cepstral coefficients . . . . . 78
6.3 Performance Metrics Comparison . . . . . 80
6.3.1 Accuracy . . . . . 80
6.3.2 Precision . . . . . 80
6.3.3 Recall . . . . . 81
6.3.4 F-score . . . . . 81
6.4 Discussion . . . . . 82

**7 Conclusion and Future Works 85**
7.1 Conclusion . . . . . 85
7.2 Future Works . . . . . 86