

CONTENTS

APPROVAL PAGE	iii
CONTENTS	vii
LIST OF TABLES	ix
LIST OF FIGURES	xi
LIST OF EQUATIONS	xiii
INTISARI	xiv
ABSTRACT	xv
I. INTRODUCTION	1
1.1 Research Background	1
1.2 Research Problem	3
1.3 Research Scope	3
1.4 Research Objective	4
1.5 Research Advantage	4
II. LITERATURE REVIEW	5
III. BASIC THEORY	9
3.1 Ensemble Learning	9
3.2 Stacked Generalization	9
3.3 Support Vector Machine	10
3.4 K-Nearest Neighbors	12
3.5 Platt Scaling	14
3.6 Confusion Matrix	14
3.7 Accuracy	15
3.8 Recall	15
3.9 Precision	16
3.10 F-Score	16
3.11 Receiving Operating Characteristics - Area Under Curve	16
IV. RESEARCH METHODOLOGY	17
4.1 Research Description	17
4.2 Research Phase	17

4.2.1 Literature Study	18
4.2.2 System Design	18
4.2.3 Data Acquisition	19
4.2.4 Data Pre-processing	21
4.2.5 SVM-KNN Ensemble Model	22
4.2.6 Level-0 Layer	23
4.2.7 Level-1 Layer	24
4.2.8 Performance Evaluation	25
V. IMPLEMENTATION	26
5.1 Import Necessary Libraries	26
5.2 Load Dataset	27
5.3 Dataset Cleaning	28
5.4 Dataset Description	29
5.5 Dataset Splitting for Base Learner and Meta Learner	30
5.6 K-Nearest Neighbors and Support Vector Machine Base Learner	32
5.7 Prediction Result	32
5.8 Probabilities Output	34
5.9 Probabilities Assignment	35
5.10 Meta Learner	36
5.11 Evaluation	37
VI. RESULTS AND DISCUSSIONS	38
6.1 Dataset Proceedings Result	38
6.2 Jakarta Climate Dataset Training and Prediction Result	41
6.3 Kerala Rainfall Dataset Training and Prediction Result	45
6.4 Data Splitting Technique Evaluation	47
6.5 Non-Splitting Technique Evaluation	53
6.6 Final Evaluation	58
V. CONCLUSIONS AND FUTURE WORKS	60
REFERENCES	61