機械学習手法と画像処理の高度応用に関する研究：交通、天気、パターン認識、ノイズフィルタリング

Form 3

Abstract

Title: **A study on advanced applications of machine learning approach and Image processing: traffic, weather, pattern recognition, and noise filtering**

Machine learning is a set of algorithms for learning concepts and finding structural patterns in data. With automatically improving experience through those learnings, estimating future events could be made successfully. Another key thing to remember is that classification, regression, clustering, density estimation, and dimensionality reduction are the core tasks of machine learning field. Thus, machine learning has been taking a vital role in the field of medical diagnosis, financial, weather, traffic jam, speech recognition, etc. Advanced algorithms for pattern classification, rainfall level prediction, traffic jam and noise filtering for rainfall radar image have been developed in this thesis.

Estimation and analysis of traffic jams plays a vital role in an intelligent transportation system and advances safety in the transportation system as well as mobility and optimization of environmental impact. For these reasons, this dissertation addresses a multi-threaded machine learning system using multisensory fusion and a hybrid approach, which stands on accuracy comparison of three prediction models: multinomial logistic regression, decision trees, and support vector machine (SVM), for traffic flow estimation.

Classification is a process of discovering and categorizing objects from large data storage that have similar characteristics, properties, and patterns. One of the most widely used classification methods in machine learning is the k-nearest neighbors (k-NN) algorithm. The unique property of k-NN that appeals to researchers is its simplicity, so it can be applied successfully over a wide field. However, according to measurement of the performance of an algorithm based on three considerations (simplicity, processing time, and prediction power), the k-NN algorithm lacks high-speed computation, high robustness and maintenance of high accuracy for different k values. To overcome these issues, this dissertation introduces plurality rule-based density and correlation coefficient-based clustering for k-NN (PRD-kNN), dual-kNN, regional distance-based kNN classification (RD-kNN), and kd-tree-based dual-kNN and confirms that the effectiveness of those algorithms by comparing with normal k-NN, density peaks clustering based on k-NN and principal component analysis (DPC-KNN-PCA), logistic regression, and neural network, and conducting experiments on real datasets from UCI machine learning repository, and rainfall radar images from the WITH small-dish aviation radar installed on the rooftop of Information Engineering, University of the Ryukyus.

As we know, machine learning algorithms are powerful tools for a variety of application domains giving widely divergent dimensions such as reliable, precision, robustness, high speed solution, etc. Likewise, the other critical dimension that a well-designed learning algorithm should occupy is a high strength of unpredictable and phenomenal noise. For this critical dimension, this research intends to investigate the attribute and class (mislabeled) noise-tolerant level of dual-kNN by injecting different noise levels. The empirical experimentations describes that dual-kNN has a higher attribute and class noise-resistant level than normal k-NN, PRD-kNN, logistic regression, and neural network.

It is never possible to able to collect a perfect real dataset due to data corruption because of sensor or acquisition devices, and data transmission. The data corruption constantly forces the learning algorithms struggle with prediction or classification works and face performance degradation in terms of mainly prediction accuracy. Thus, this dissertation presents a noise filtering algorithm, named as an adaptive morphological operation for high-performance weather image processing. The experimental results based on 2011, 2013, 2015, and 2016 datasets, confirms that the adaptive approach is more efficient than the conventional morphological operations.

Name Swe Swe Aung