

นิเวศน์เดวีร์กเป็นแนวทางหนึ่งที่ถูกพัฒนาขึ้นเพื่อใช้ในการจำแนกข้อมูล แต่ทั้งนี้สิ่งที่เป็นอุปสรรคต่อการพัฒนาความสามารถในการจำแนกข้อมูลคือ จำนวนข้อมูลปรับสอนที่มีจำกัดทำให้นิเวศน์เดวีร์กไม่สามารถเรียนรู้ลักษณะของข้อมูลได้เพียงพอที่จะสามารถจำแนกข้อมูลให้ถูกต้องในระดับที่น่าเชื่อถือได้ โครงการวิจัยนี้เป็นการศึกษาแนวทางในการแก้ปัญหานี้จำนวนข้อมูลปรับสอนมีจำกัด

ข้อมูลที่ใช้เป็นกรณีศึกษาในโครงการวิจัยนี้คือข้อมูล Control Chart Pattern Recognition และข้อมูลผู้ป่วยโรคหัวใจ จากสถานะที่มีข้อมูลเพียงพอนั้นชุดข้อมูลทั้งสองได้ถูกลดจำนวนลงในสัดส่วนที่ไม่เท่ากันในแต่ละคลาส ความถูกต้องในการจำแนกข้อมูลจากชุดข้อมูลที่จำกัดอยู่ในระดับที่ดีที่สุดคือ 83.33% และ 80.00% ตามลำดับ

โครงการวิจัยนี้นำเสนอผลจากการประยุกต์แนวทาง Samples Duplication, Cross-validation, Leave-one-out cross validation และ Synergistic Neural Networks เพื่อทำการปรับปรุงประสิทธิภาพการจำแนกข้อมูลที่มีจำนวนจำกัดดังกล่าว ผลของการวิจัยพบว่าแนวทาง Samples Duplication ที่ใช้ควบคู่ไปกับแนวทาง Synergistic Neural Networks สามารถเพิ่มประสิทธิภาพการจำแนกข้อมูล Control Chart Pattern Recognition ได้ดีที่สุดและการประยุกต์ใช้แนวทาง Samples Duplication กับข้อมูลผู้ป่วยโรคหัวใจสามารถเพิ่มประสิทธิภาพการจำแนกได้ดีที่สุดโดยให้ระดับค่าความถูกต้องในการจำแนกข้อมูลที่ 93.33% และ 86.67% ตามลำดับ

Neural networks is one of many techniques used in classification. A situation which prevents good performance in this task is the limited number of available samples, since neural networks is not able to learn efficiently from the training samples in order to reach a satisfactory performance. This work presents techniques which could be used to improve performance of the classification under such constrain.

The data used for the work were time series in control chart pattern recognition and heart disease data. From original sources with sufficient number of samples, the sample sets were then reduced considerably and unproportionally in each class. These hampered the performance to no more than 83.33% and 80.00% accuracy respectively.

Techniques known as samples duplication, cross-validation , leave-one-out cross validation and synergistic neural networks were then utilized in order to compensate for the limited samples. The results reveals that samples duplication together with synergy of neural networks yielded the highest improvement for the classification of control chart pattern recognition data while samples duplication did so for the heart disease data with the accuracy level of 93.33% and 86.67% respectively.

Keywords : classification / control chart pattern recognition / cross-validation / heart disease data / leave-one-out cross validation / limited number of samples / neural networks / samples duplication / synergistic neural networks