

ชื่อเรื่องวิทยานิพนธ์	การประยุกต์เทคนิคการวิเคราะห์หลายตัวแปรกับข้อมูลดีเอ็นเอไมโครอาร์เรย์
ผู้เขียน	นาย จำลอง คำผาเชื้อ
ปริญญา	วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผู้ช่วยศาสตราจารย์ ดร.จิรยุทธ ไชยจารุณิช

บทคัดย่อ

งานวิจัยมีวัตถุประสงค์เพื่อประยุกต์เทคนิคการวิเคราะห์หลายตัวแปรกับข้อมูลดีเอ็นเอไมโครอาร์เรย์ที่ลักษณะของข้อมูลดังกล่าวประกอบไปด้วยตัวแปรจำนวนมาก ทั้งนี้ในการวิเคราะห์ได้ใช้วิธีการต่างๆ ได้แก่ การวิเคราะห์องค์ประกอบหลัก การวิเคราะห์ปัจจัย การวิเคราะห์การจำแนกประเภท และการวิเคราะห์การถดถอยแบบโลจิสติก โดยใช้ การวิเคราะห์องค์ประกอบหลักในการลดทอนข้อมูล เพื่อนำเสนอค่าการแสดงผลของยีน ในกรณีที่มีตัวแปรจำนวนมาก ให้อยู่ในมิติที่สามารถรับรู้ได้ ผลของการนำเสนอด้วยวิธีนี้ทำให้มองเห็นลักษณะการกระจายตัวของยีนและตัวอย่างข้อมูลที่ศึกษาได้ง่ายขึ้น ในมุมมองขององค์ประกอบหลักพบว่าทำให้มองเห็นข้อแตกต่างของกลุ่มข้อมูลดังกล่าวได้อย่างชัดเจน การวิเคราะห์ปัจจัยเพื่อหาโครงสร้างของตัวแปร ทั้งในด้านยีนและเงื่อนไขการทดลอง พบว่าวิธีการดังกล่าวช่วยในการลดทอนข้อมูล การนำเสนอข้อมูล การอธิบายความหมายของข้อมูลและการจัดกลุ่มข้อมูล ได้อย่างมีประสิทธิภาพ การวิเคราะห์การจำแนกประเภทและการวิเคราะห์การถดถอยแบบโลจิสติก วิธีการทั้งสองมีจุดประสงค์เพื่อที่จะจำแนกกลุ่มของข้อมูล จากงานวิจัย พบว่าทั้งสองวิธีการสามารถจำแนกกลุ่มของข้อมูลได้ดีในระดับหนึ่ง แม้ว่าผลการวิเคราะห์จากเทคนิคการวิเคราะห์หลายตัวแปรเหล่านี้ ยังไม่ได้ยืนยันโดยข้อสรุปเชิงชีววิทยา หากแต่งานวิจัยได้แสดงให้เห็นถึงการประยุกต์วิธีการดังกล่าวกับข้อมูลดีเอ็นเอไมโครอาร์เรย์ในหลายลักษณะ ซึ่งมีประโยชน์ต่อการศึกษาวิจัยทางด้านชีวสารสนเทศ ในขั้นสูงต่อไป

Thesis Title	The Application on Multivariate Analysis Techniques to DNA Microarray Data
Author	Mr. Jamlong Khamphachua
Degree	Master of Science (Computer Science)
Thesis Advisor	Assistant Professor Dr. Jeerayut Chaijaruwanich

ABSTRACT

The objective of this research is to apply the multivariate analysis approach to analyzing high dimensional DNA microarray data. Four methods of this technique such as principal component analysis, factor analysis, discriminant analysis and logistic regression analysis are applied to various DNA microarray data sets. The first one is used in data reduction and representing the distribution of gene expression data on fewer dimensions. In the principal component space, the different groups of data can be visualized explicitly. The factor analysis is applied to detect the structure of variables, i.e. genes and experiment conditions. It effectively helped in data reduction, data representation, data interpretation and data clustering. The last two methods: the discriminant analysis and the logistic regression analysis are used for data classification. The experiment results show that both methods have efficiency in discriminating and predicting the class of data. Nevertheless, our results are not yet confirmed by biological evidences, this research illustrates the various applications of multivariate analysis to DNA microarray data which will be significant for advanced bioinformatics research.