

# Guest Editors' Introduction: *Special Issue on Data Mining*

EDISI khusus IJ ini memilih datamining sebagai topik utama. Datamining merupakan salah satu topik menarik dalam bidang komputasi yang akhir-akhir ini mendapatkan perhatian yang makin meningkat, baik dari kalangan akademisi maupun praktisi di berbagai sektor. Hal ini tidak terlepas dari ketersediaan data yang berskala besar pada berbagai bidang, sehingga menuntut pengembangan teknik-teknik komputasi untuk mengekstrak informasi signifikan dari lautan data tsb. Ketersediaan data dalam skala besar ini dijumpai di berbagai bidang, antara lain bioteknologi, meteorologi, akademik dan finansial. Terlebih lagi dewasa ini penelitian lintas bidang seolah sudah menjadi keharusan, karena pemecahan masalah yang demikian kompleks tidak akan tercapai tanpa adanya kerjasama berbagai disiplin ilmu. Hal ini menyebabkan data yang diperoleh tidak hanya berskala besar, tetapi juga semakin kompleks dan heterogen. Tantangan-tantangan inilah yang menyebabkan datamining menjadi salah satu tren teknologi terkini dalam bidang komputasi. Datamining diharapkan mampu menemukan informasi yang tersuruk dalam lautan data itu untuk dimanfaatkan dalam berbagai aplikasi, baik di dunia kedokteran, marketing maupun bidang yang lain.

Di Indonesia, minat terhadap datamining terlihat makin meningkat, sebagaimana tercermin pada diskusi di berbagai komunitas maya. Berbagai perguruan tinggi telah memasukkan datamining dalam kurikulumnya, disamping matakuliah yang secara konten berkaitan erat seperti statistika, pengenalan pola dan machine learning. Pelatihan datamining pun telah dilakukan di berbagai instansi, seperti perbankan misalnya. Kebutuhan akan praktisi TI yang menguasai teknologi datamining pun mulai sering terlihat dari proses perekrutan pegawai, misalnya perekrutan pegawai perbankan di Indonesia. Hal lain yang menggembirakan adalah mulai ditemuinya buku-buku teks datamining berbahasa Indonesia, sehingga memudahkan para pelajar, mahasiswa dan peneliti untuk melakukan studi dan penelitian di bidang ini. Hal-hal ini yang melatarbelakangi dipilihnya topik datamining dalam edisi khusus IJ kali ini.

Dari berbagai makalah yang diterima, dilakukan proses review oleh para pakar, dan akhirnya terpilih 7 makalah untuk dimuat pada edisi kali ini. Paper pertama memperkenalkan algoritma Enhanced SMART-TV (SMall Absolute difference of ToTal Variation) yang merupakan varian dari nearest neighbor yang memiliki kelebihan pada sisi kecepatan, dibandingkan dengan pendekatan nearest neighbor konvensional. Paper kedua mengambil topik Business

Intelligence untuk perancangan strategi, dengan studi kasus pada Arab International University. Paper ketiga mengupas teknik feature selection untuk aplikasi pada data skala besar. Paper keempat merupakan aplikasi datamining untuk informasi tekstual, yaitu mendeteksi originality sebuah paper terhadap paper yang telah dipublikasikan sebelumnya. Dengan demikian, bisa diketahui apakah seorang peneliti memang benar-benar telah berhasil mengembangkan ide penelitiannya, ataukah sekedar mengulang ide yang telah dipresentasikan yang bukan lain adalah bentuk plagiarisme. Paper kelima adalah pemenang kompetisi datamining nasional bagi mahasiswa S1 yang diselenggarakan oleh Dikti pada GeMasTIK II tahun 2009, bertemakan prediksi masa studi sarjana berdasarkan data-data akademis. Paper keenam mengajukan teknik baru untuk membangun sistem generasi navigasi berdasar konten secara otomatis. Paper ketujuh menengahkan aplikasi datamining pada kesehatan, yaitu dalam menganalisa data diabetes. Teknik datamining yang dibahas dalam ketujuh tema tersebut cukup beragam, dengan target aplikasi pada bidang akademik dan kesehatan.

Edisi khusus ini terbit dengan dukungan para reviewer yang merupakan pakar data mining Indonesia. Untuk itu kami ucapkan terimakasih dan penghargaan yang tinggi kepada beberapa kolega yang telah melakukan review; Dr. Budi Santosa, Dr. Yudho Giri Sucahyo, Dr. G.A. Putri Saptawati, Prof. The Houw Liong, Dr. Yudhi Agusta, Dr. Iko Pramudiono, dan terutama kepada Prof David Taniar yang juga telah mendorong adanya edisi khusus ini. Dan tentunya kepada editor IJ Dr. Thomas Harjono yang telah mengawal dan mendukung dari awal hingga akhir edisi ini.

Kami berharap agar paper yang diterbitkan pada Edisi Khusus kali ini dapat bermanfaat bagi penelitian dan pengembangan aplikasi datamining pada berbagai sektor di tanah air. Semoga pada masa mendatang semakin banyak tulisan berkaitan dengan datamining yang dapat dimuat pada IJ.

---

This Special Issue of the IJ focuses on data mining. Data mining is an interesting topic within the broad field of computation that has received increased attention recently, not only from academics but also from practitioners in various sectors. This is not unrelated to the availability of large scale data within different areas, and which therefore has necessitated the development of new computational techniques to extract significant information from this ocean of data. The availability of large scale data can be found in various areas, including biotechnology, meteorology, the academic fields and the financial sector. Furthermore, in recent times multidisciplinary research has in reality become

The Guest Editors can be reached at the following email addresses. Dr. Anto Satriyo Nugroho is at [asnugroho@inn.bppt.go.id](mailto:asnugroho@inn.bppt.go.id), while Moch Arif Bijaksana is at [mab@ittelkom.ac.id](mailto:mab@ittelkom.ac.id).

unavoidable (and even mandatory) due to the fact that solving many of these complex problems requires collaboration across various disciplines of study. This has resulted in data not only being large scale, but also more complex and heterogeneous. It is these challenges that have made data mining a new technological trend of interest within the field of computation. The expectation is that data mining can be deployed to find useful information buried within the vast ocean of data, which in turn can be beneficial to diverse areas of application such as medicine, marketing and others.

In Indonesia recently there has been an increase in interest in data mining, as reflected by the discussion traffic on the Internet. A number of universities and higher education institutions have introduced the subject of data mining into their curriculum as an addition to subjects whose contents are already closely related to data mining, such as statistics, pattern recognition and machine learning. Training in data mining has also been provided by a number of institutions, such as those in the banking sector. The increasing demand for IT practitioners who are familiar with data mining technologies is also evident from increasing recruitments seeking these practitioners, such as in the area of banking. Most pleasing also is the increasing availability of textbooks on data mining written in Bahasa Indonesia, which allows a broader range of students and researchers alike to focus on the study of data mining and conduct research in this area. It is all these factors that have provided the backdrop for this IJ Special Issue on Data Mining.

This Special Issues carries seven (7) papers which have been reviewed by experts in this field. The first paper introduces the Enhanced SMART-TV algorithm (SMall Absolute diffeRence of ToTal Variation), which is a variation of the nearest neighbor approach. However, this algorithm variation provides an increase in speed compared to the conventional nearest neighbor approach. The second paper discusses on the role of business intelligence in achieving the business strategy of an organization, using the Arab International University as its case-study. The third paper focuses on a feature selection technique for large scale data. The fourth paper uses data mining on contextual information for the purpose of detecting the originality of the contents of a new paper as compared to existing published papers. The aim here is to determine whether a researcher has truly expanded a given research idea, or if he or she has simply repeated the same idea using a different presentation, something considered by many to be equivalent to plagiarism. The fifth paper is the winner of the Indonesian national data mining competition for S1 (undergraduate) students held by the DIKTI organization (Directorate General for Higher Education) at the GeMasTIK II event in 2009, which is a student event focusing on information and communications technologies. The theme of this fifth paper is the predictive techniques that can be used to predict a student's performance based on their previous academic data. The sixth paper is on a navigation-generating system based on adaptive content, while the seventh paper discusses the application of data mining within the field of healthcare, for the purpose of analyzing diabetes related data.

All in all the data mining techniques discussed within these seven papers cover a broad range, with the target area of application primarily being those of the academic field and healthcare.

This Special Issue on Data Mining received tremendous support from reviewers who are experts in the field of data mining in Indonesia. We would like to thank and indicate our deepest appreciation for these colleagues who did the reviews: Dr. Budi Santosa, Dr. Yudho Giri Sucahyo, Dr. G.A. Putri Saptawati, Prof. The Houw Liong, Dr. Yudhi Agusta, and Dr. Iko Pramudiono. Special appreciation goes to Prof. David Taniar who came-up with the idea of an IJ Special Issue and strongly supported its creation. We also thank the IJ Editor Dr. Thomas Hardjono who has provided editorial guidance and support from the beginning until the completion of this Special Issue.

We hope the papers appearing in this IJ Special Issue on Data Mining will be of benefit to research and development on data mining applications within the various sectors in Indonesia. We hope other papers related to data mining can appear in the IJ in the future.

Anto Satriyo Nugroho  
Moch Arif Bijaksana



**Dr. Anto Satriyo Nugroho** is a researcher working for Agency for the Assessment & Application of Technolog (BPPT), Indonesia. He received his B.Eng degree in 1995, M.Eng in 2000 and Dr.Eng degree in 2003, all in Electrical & Computer Engineering from Nagoya Institute of Technology, Japan. From 2003 to 2007 he worked for Chukyo University as a Visiting Professor in the School of Cognitive and Computer Science (2003-2004) and in the School of Life System Science & Technology (2004-2007). He received the First Prize award of the '99 Fog Forecasting Contest conducted by the Neurocomputing Technical Group of The Institute of Electronics, Information and Communication Engineers (IEICE), Japan, and also the best paper award in the Konferensi Nasional Sistem & Informatika 2008 at Inna Sindhu Beach Hotel, Sanur Bali. His research interests are in the areas of pattern recognition, computational biology and data mining. Since 2007 he has been serving as the vice president of the Indonesian Society for Softcomputing. Dr. Nugroho is a member of the Institute of Electrical & Electronics Engineers (IEEE).



**Moch Arif Bijaksana** is a lecturer working for Institut Teknologi Telkom (IT Telkom), Indonesia. He received his B.Eng degree in 1991 from Gadjah Mada University and M.Tech in 1998 from RMIT University. He is an active member of Indonesian Data Mining Community. His main research interest is text mining.