

TUGAS AKHIR

**PENERAPAN *DATA MINING* DENGAN ALGORITMA
CLASSIFICATION BASED ON ASSOCIATION (CBA)
UNTUK KLASIFIKASI RISIKO PEMBERIAN KREDIT
(Studi Kasus : PT. Pupuk Sriwijaya Palembang)**

Jenjang Strata 1



OLEH

FIRZA FAUZIYYAH

NIM 09031181419016

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVESITAS SRIWIJAYA**

2018

LEMBAR PENGESAHAN

**PENERAPAN *DATA MINING* DENGAN ALGORITMA
CLASSIFICATION BASED ON ASSOCIATION (CBA)
UNTUK KLASIFIKASI RISIKO PEMBERIAN KREDIT
(Studi Kasus : PT. Pupuk Sriwidjaja Palembang)**

SKRIPSI

Sebagai salah satu syarat untuk penyelesaian
studi di Program Studi Sistem Informasi
Jenjang S1

Oleh :

FIRZA FAUZIYYAH
(09031181419016)

Palembang, Maret 2018

Pembimbing I,



Ken Ditha Tania, M.Kom
NIP. 198507182012122003

Pembimbing II,



Allsela Meiriza, M.T
NIP. 198305132015012201

Mengetahui,

Ketua Jurusan Sistem Informasi,



Endang Lestari Ruskan, M.T.
NIP. 197811172006042001

HALAMAN PERSETUJUAN


Telah lulus dan diuji pada :


Hari : Sabtu

Tanggal : 24 Maret 2018

Tim Penguji :

1. Pembimbing I : Ken Ditha Tania, M.Kom
2. Pembimbing II : Allsela Meiriza, M.T
3. Ketua : Pacu Putra Suarli, M.CS
4. Anggota I : Yadi Utama, M.Kom.
5. Anggota II : Ali Ibrahim, M.T.







**Mengetahui,
Ketua Jurusan Sistem Informasi,**



**Endang Lestari R, M.T.
NIP. 197811172006042001**

HALAMAN PERNYATAAN

Nama : Firza Fauziyyah
NIM : 09031181419016
Program Studi : Sistem Informasi
Judul Skripsi : Penerapan *Data Mining* dengan
Algoritma Classification Based on Association (CBA) Untuk Klasifikasi Risiko
Kredit (Studi Kasus: PT. Pupuk Sriwidjaja Palembang)
Hasil Pengecekan Software Ithenticate/Turnitin : 10 %

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan /plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikianlah, Pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, Maret 2018




(Firza Fauziyyah)

NIM. 09031181419016

MOTTO DAN PERSEMBAHAN

MOTTO :

*Sometimes you find out what you are supposed to be doing by doing the things
you are not supposed to do.*

(Oprah Winfrey)

Tugas Akhir ini saya persembahkan untuk :

- *Ayah, Bunda, dan Adik tercinta*
- *Dosen Pembimbing beserta Dosen Penguji*
- *Sahabat-sahabat saya*
- *Teman-teman seperjuangan di Sistem Informasi
2014*
- *Almamaterku Universitas Sriwijaya*

**PENERAPAN *DATA MINING* DENGAN ALGORITMA
CLASSIFICATION BASED ON ASSOCIATION (CBA)
UNTUK KLASIFIKASI RISIKO PEMBERIAN KREDIT
(Studi Kasus : PT. Pupuk Sriwidjaja Palembang)**

Oleh

Firza Fauziyyah 09031181419016

ABSTRAK

Untuk mencegah terjadinya kredit bermasalah pada pinjaman dana untuk usaha mikro yang disebabkan oleh faktor internal, diperlukan suatu sistem yang dapat memprediksi risiko kredit calon peminjam dana dengan lebih akurat dan objektif. Salah satunya adalah dengan memanfaatkan teknik *data mining* menggunakan algoritma *Classification Based on Association (CBA)* yang dapat menghasilkan model untuk memprediksi risiko kredit dengan mengklasifikasi data calon peminjam dana ke dalam kelas lancar atau tidak lancar. Proses *data mining* dilakukan dengan menggunakan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. Hasil prediksi dengan menggunakan model yang dihasilkan algoritma CBA dengan bahasa R kemudian diuji tingkat akurasi. Model dengan tingkat akurasi tertinggi kemudian diimplementasikan ke dalam sebuah sistem aplikasi berbasis *web*.

Kata Kunci : *Data mining, classification based on association, CBA, risiko pemberian kredit, klasifikasi, CRISP-DM.*

**IMPLEMENTATION OF DATA MINING USING
CLASSIFICATION BASED ON ASSOCIATION (CBA) ALGORITHM
TO CLASSIFY CREDIT RISKS
(CASE STUDY : PT. PUPUK SRIWIDJAJA PALEMBANG)**

By

Firza Fauziyyah 09031181419016

ABSTRACT

In order to prevent loan defaults on micro enterprises which caused by internal factors, it needs a system that able to predict credit risk of loan applicants more accurately and objectively. One of the ways to do so is by utilizing data mining technique using Classification Based on Association (CBA) algorithm that able to build a model to predict credit risks by classifying loan applicants data into “good” or “delinquent” class. The data mining process was done by using Cross-Industry Standard Process for Data Mining (CRISP-DM) method. The accuracy level of the resulting model using the CBA algorithm with R language then was tested. A model with the highest accuracy level was then implemented into a web based application system.

Keywords : Data mining, classification based on association, CBA, credit risk, classification, CRISP-DM.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat, taufik serta hidayah-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul “Penerapan *Data Mining* dengan Algoritma *Classification Based on Association (CBA)* Untuk Klasifikasi Risiko Pemberian Kredit (Studi Kasus: PT. Pupuk Sriwidjaja Palembang)” dengan tanpa adanya halangan yang berarti.

Laporan tugas akhir ini merupakan salah satu syarat wajib yang harus ditempuh untuk dapat menyelesaikan studi di Program Studi Sistem Informasi Jenjang Sarjana. Dalam penyusunan laporan tugas akhir ini penulis banyak mendapatkan bantuan dari berbagai pihak, oleh sebab itu penulis ingin mengungkapkan rasa terima kasih kepada :

1. Bapak Prof. Dr. Ir. H. Anis Saggaff ,MSCE selaku Rektor Universitas Sriwijaya
2. Bapak Jaidan Jauhari, M.T selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya
3. Ibu Endang Lestari Ruskan, M.T selaku Ketua Jurusan Sistem Informasi Fasilkom Universitas Sriwijaya
4. Bapak Ahmad Rifai, S.T., M.T selaku Dosen Pembimbing Akademik penulis
5. Ibu Ken Ditha Tania, M.Kom dan Ibu Allsela Meiriza, M.T selaku Dosen Pembimbing Tugas Akhir penulis
6. Bapak Yadi Utama, M.Kom dan Bapak Ali Ibrahim, M.T selaku Dosen Penguji Tugas Akhir penulis
7. Departemen Program Kemitraan dan Bina Lingkungan PT. Pupuk Sriwidjaja Palembang yang telah memberi kesempatan kepada penulis untuk melaksanakan kegiatan penelitian guna penyusunan laporan tugas akhir ini

8. Seluruh staf jurusan Sistem Informasi dan akademik Fakultas Ilmu Komputer Universitas Sriwijaya yang telah memberikan kemudahan selama pengurusan administrasi
9. Kedua orang tua, adik, serta keluarga besar penulis yang telah memberikan dukungan dan semangat terbesar dalam proses penyusunan laporan tugas akhir ini
10. Sahabat – sahabat terbaik penulis Ade Riska, Ayu Zahrah, Daisa Rani, Fairuz, Maretha Sari, Rahmi Mahmuda dan Riri Yunita yang telah menemani penulis selama menjalani masa studi
11. Teman - teman Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya Angkatan 2014
12. Semua pihak yang telah memberikan dukungan kepada penulis yang tidak dapat penulis sebutkan satu – persatu.

Penulis menyadari dalam pelaksanaan maupun pembuatan laporan ini penulis tidaklah sempurna dan masih memiliki banyak kekurangan. Apabila nantinya terdapat kekeliruan dalam penulisan laporan tugas akhir ini, penulis sangat mengharapkan kritik dan saran yang membangun.

Akhir kata, semoga laporan ini dapat bermanfaat bagi penulis dan orang banyak.

Indralaya, Maret 2018

Penulis,

Firza Fauziyyah

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
ABSTRAK	iii
ABSTRACT	iv
DAFTAR ISI	v
DAFTAR GAMBAR	ix
DAFTAR TABEL	xi
BAB I. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Tujuan	3
1.3. Manfaat	3
1.4. Batasan Masalah.....	4
BAB II. TINJAUAN PUSTAKA	5
2.1. Kajian Pustaka	5
2.2. Profil Perusahaan.....	7
2.2.1. Sejarah PT. Pupuk Sriwidjaja	7
2.2.2. Visi, Misi dan Nilai-nilai Perusahaan	9
2.2.3. Struktur Organisasi PT. Pupuk Sriwidjaja Palembang	10
2.2.4. Struktur Organisasi Divisi Sekretariat Perusahaan & Tata Kelola	11
2.2.5. Struktur Organisasi Departemen PKBL.....	11
2.2.6. Fungsi dan Tugas Pokok Jabatan Departemen PKBL	12
2.2.7. Program Kemitraan dan Bina Lingkungan	17
2.2.8. Program Kemitraan	18

2.3.	<i>Data Mining</i>	21
2.3.1.	Pengertian <i>Data Mining</i>	21
2.3.2.	<i>Knowledge Discovery Process</i>	23
2.3.3.	Proses <i>Data Mining</i>	24
2.3.4.	Klasifikasi	25
2.3.5.	<i>Classification Based on Association (CBA)</i>	26
2.3.6.	Bahasa R	31
2.3.7.	Jupyter Notebook	33
2.4.	Pemodelan Data	33
2.4.1.	<i>Entity Relationship Diagram</i>	34
2.4.2.	<i>Data Flow Diagram</i>	36
2.5.	PHP	37
2.6.	MySQL	38
2.7.	Web.....	38
BAB III.	METODOLOGI PENELITIAN	40
3.1.	Subjek dan Objek Penelitian	40
3.2.	Lokasi Penelitian	40
3.3.	Teknik Pengumpulan Data.....	40
3.3.1.	Jenis Data	40
3.3.2.	Sumber Data.....	41
3.3.3.	Metode Pengumpulan Data	41
3.4.	Metode <i>Data Mining</i>	42
3.4.1.	Fase Pemahaman Bisnis.....	43
3.4.2.	Fase Pemahaman Data	43
3.4.3.	Fase Persiapan Data	43
3.4.4.	Fase Pemodelan	43

3.4.5. Fase Evaluasi	44
3.4.6. Fase Penyebaran.....	44
3.5. Metode Pengembangan Perangkat Lunak	45
3.6. Kerangka Berpikir	48
BAB IV. ANALISIS SISTEM.....	49
4.1. Fase Pemahaman Bisnis.....	49
4.1.1. Analisis Kebutuhan Fungsional	49
4.1.2. Analisis Kebutuhan Non Fungsional	50
4.1.3. Alur Kerja Sistem	51
4.2. Fase Pemahaman Data	51
4.3. Fase Persiapan Data	54
4.3.1. Pemilihan Data.....	54
4.3.2. Pembersihan Data	55
4.3.3. Transformasi Data.....	55
4.4. Fase Pemodelan	56
4.5. Fase Evaluasi	62
BAB V. RANCANGAN SISTEM	64
5.1. Rancangan Arus Data	64
5.1.1. DFD Level 0	64
5.1.2. DFD Level 1	65
5.2. PDFD	67
5.3. Rancangan Basis Data.....	70
5.3.1. ERD.....	70
5.3.2. Struktur Tabel	71
5.3.3. Rancangan Skema Basis Data.....	74
5.4. Rancangan Antar Muka	75

5.4.1. Rancangan Halaman <i>Login</i>	75
5.4.2. Rancangan Halaman <i>Home</i>	76
5.4.3. Rancangan Halaman Klasifikasi	77
5.4.4. Rancangan Halaman Hasil Klasifikasi.....	78
BAB VI. HASIL DAN PEMBAHASAN	79
6.1. Hasil	79
6.2. Pembahasan.....	80
6.2.1. Halaman <i>Login</i>	80
6.2.2. Halaman <i>Home</i>	82
6.2.3. Halaman Klasifikasi.....	83
6.2.4. Halaman Hasil Klasifikasi	85
6.3. Pengujian Sistem.....	86
BAB VII. KESIMPULAN DAN SARAN	93
7.1. Kesimpulan	93
7.2. Saran	94
DAFTAR PUSTAKA	95
LAMPIRAN	98

DAFTAR GAMBAR

Gambar 2.1	Struktur Organisasi PT. Pusri.....	10
Gambar 2.2	Struktur Organisasi Divisi Sekretariat Perusahaan & Tata Kelola .	11
Gambar 2.3	Struktur Organisasi Departemen PKBL.....	11
Gambar 2.4	Mekanisme Penyaluran Program Kemitraan di BUMN	20
Gambar 2.5	<i>Knowledge Discovery Process</i>	23
Gambar 2.6	(a) aturan <i>IF-THEN</i> , (b) <i>decision tree</i> , (c) <i>neural network</i>	26
Gambar 2.7	Algoritma CBA-RG menurut Liu et al.....	28
Gambar 2.8	Algoritma M2 Tahap 1	29
Gambar 2.9	Algoritma M2 Tahap 2.....	30
Gambar 2.10	Algoritma M2 Tahap 3.....	30
Gambar 2.11	Algoritma M2 Tahap 4.....	31
Gambar 2.12	Contoh ERD	36
Gambar 3.1	6 Fase Metode CRISP-DM	42
Gambar 3.2	5 Fase Metode Waterfall	46
Gambar 3.3	Kerangka Berpikir.....	48
Gambar 4.1	Alur Kerja Sistem.....	51
Gambar 4.2	Data Mentah	42
Gambar 4.3	Tahapan <i>Data Training</i> 1	57
Gambar 4.4	Tahapan <i>Data Training</i> 2.....	58
Gambar 4.5	Tahapan <i>Data Training</i> 3.....	58
Gambar 4.6	Tahapan <i>Data Training</i> 4.....	59
Gambar 4.7	Tahapan <i>Data Training</i> 5.....	59
Gambar 4.8	Tahapan <i>Data Training</i> 6.....	60
Gambar 4.9	Tahapan <i>Data Testing</i> 1	61

Gambar 4.10 Tahapan <i>Data Testing</i> 2	61
Gambar 4.11 Hasil Prediksi	62
Gambar 5.1 DFD Level 0.....	64
Gambar 5.2 DFD Level 1.....	65
Gambar 5.3 PDFD	66
Gambar 5.4 ERD.....	70
Gambar 5.5 Rancangan Skema Basis Data.....	75
Gambar 5.6 Rancangan Halaman <i>Login</i>	75
Gambar 5.7 Rancangan Halaman <i>Home</i>	76
Gambar 5.8 Rancangan Halaman Klasifikasi	77
Gambar 5.9 Rancangan Halaman Hasil Klasifikasi.....	78
Gambar 6.1 Halaman <i>Login</i>	80
Gambar 6.2 Notifikasi <i>Login</i> Gagal.....	81
Gambar 6.3 Halaman <i>Home</i>	82
Gambar 6.4 Halaman Klasifikasi	83
Gambar 6.5 <i>Modal</i> Tambah Data.....	83
Gambar 6.6 Notifikasi Data Belum Lengkap.....	84
Gambar 6.7 Halaman Hasil Klasifikasi.....	85

DAFTAR TABEL

Tabel 2.1	Tabel Simbol Komponen ERD	35
Tabel 2.2	Tabel Simbol Elemen DFD.....	37
Tabel 4.1	Kebutuhan Fungsional	49
Tabel 4.2	Kebutuhan Non Fungsional	50
Tabel 4.3	Struktur Tabel Data Mentah.....	52
Tabel 4.4	Struktur <i>Dataset</i> Awal	54
Tabel 4.5	Tabel Hasil Percobaan	63
Tabel 5.1	Struktur Tabel <i>User</i>	71
Tabel 5.2	Struktur Tabel Proposal	72
Tabel 6.1	Pengujian Kebutuhan Sistem	86
Tabel 6.2	Pengujian Halaman <i>Login</i>	88
Tabel 6.3	Pengujian Halaman Klasifikasi.....	89
Tabel 6.4	Pengujian Halaman Hasil Klasifikasi	92

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sesuai Peraturan Menteri Badan Usaha Milik Negara Nomor PERMEN-05/MBU/2007 Tentang Program Kemitraan dan Bina Lingkungan Badan Usaha Milik Negara, sebagai salah satu anak perusahaan dari PT. Pupuk Indonesia Holding Company yang merupakan perusahaan Badan Usaha Milik Negara (BUMN), PT. Pupuk Sriwidjaja Palembang yang selanjutnya disebut PT. Pusri, memiliki tanggung jawab sosial (*Corporate Social Responsibility*) kepada masyarakat dan lingkungan sekitar BUMN yang diwujudkan dalam bentuk Program Kemitraan dan Bina Lingkungan (PKBL) yang terdiri dari dua program berbeda, yaitu (1) Program Kemitraan dan (2) Program Bina Lingkungan. Program Kemitraan adalah sebuah program dimana perusahaan memberikan kredit lunak dan pembinaan bagi pelaku Usaha Kecil Menengah (UKM). Sedangkan Program Bina Lingkungan adalah program berbentuk bantuan yang dapat diberikan kepada korban bencana alam, sarana ibadah, pendidikan atau pelatihan, pengembangan sarana umum dan lain-lain.

Untuk Program Kemitraan, pengajuan kredit dapat diperoleh melalui tahapan pengajuan proposal oleh calon mitra binaan, seleksi dan evaluasi oleh BUMN pembina. Dalam tahapan seleksi dan evaluasi ini prosedur dan kebijakannya ditetapkan sendiri oleh BUMN pembina dengan tujuan memilih calon mitra binaan yang potensial dan meminimalisir terjadinya kredit

bermasalah, karena adanya kredit bermasalah dapat mengganggu kondisi keuangan PT. Pusri dan berpotensi menyebabkan berkurangnya jumlah dana yang disalurkan kepada para pelaku UKM periode berikutnya.

Menurut (Addae-korankye, 2014), ada beberapa faktor penyebab terjadinya kredit bermasalah pada pinjaman usaha mikro yang disebabkan oleh pihak internal diantaranya adalah penilaian yang kurang tepat, kurangnya pengawasan, dan pemilihan penerima pinjaman yang salah. Untuk mencegah terjadinya hal ini, maka diperlukan suatu sistem terkomputerisasi yang dapat memprediksi kelancaran kredit calon mitra binaan dengan lebih akurat dan memiliki tingkat ketelitian analisis yang tinggi dan bersifat objektif. Salah satu teknik yang dapat digunakan adalah teknik *data mining*.

Menurut (J. Zaki & Meira, 2014) *data mining* terdiri dari algoritma – algoritma inti yang memungkinkan seseorang untuk memperoleh wawasan dan pengetahuan mendasar dari data yang besar. *Output* dari *data mining* yang berupa pengetahuan yang dihasilkan berdasarkan data-data mitra binaan di tahun-tahun sebelumnya inilah yang dapat digunakan untuk memprediksi kelancaran kredit calon mitra binaan di masa depan. Diantara banyak metode dalam *data mining* dapat digunakan untuk menyelesaikan permasalahan serupa, terdapat sebuah metode yang merupakan perpaduan metode klasifikasi dan asosiasi dengan algoritma *Classification Based on Association* (CBA) yang telah teruji keakuratannya dibanding beberapa algoritma lainnya (Amseke; & Winarko, 2014).

Berdasarkan latar belakang diatas, penulis tertarik untuk membuat tugas akhir dengan judul “Penerapan *Data Mining* Dengan Algoritma *Classification Based On Association* (CBA) Untuk Klasifikasi Risiko Pemberian Kredit (Studi Kasus : PT. Pupuk Sriwidjaja Palembang)”.

1.2. Tujuan

Tujuan yang ingin dicapai dalam pembuatan tugas akhir ini yaitu menerapkan teknik *data mining* dengan menggunakan algoritma *Classification Based On Association* (CBA) untuk klasifikasi risiko pemberian kredit pada Program Kemitraan dan Bina Lingkungan (PKBL) di PT. Pupuk Sriwidjaja Palembang.

1.3. Manfaat

Adapun manfaat yang diharapkan dari penelitian ini antara lain :

1. Manfaat Teoritis : Sebagai pembelajaran dalam bentuk teori dan dapat berguna sebagai referensi bagi penelitian sejenis di masa mendatang.
2. Manfaat Praktis : Hasil penelitian dapat membantu PT. Pupuk Sriwidjaja Palembang, khususnya Departemen Program Kemitraan dan Bina Lingkungan dalam menyeleksi calon mitra binaan untuk ditetapkan sebagai penerima kredit.

1.4. Batasan Masalah

Agar fokus dari penelitian ini tidak melebar ke pembahasan yang lebih luas, maka penulis menentukan batasan masalah, yaitu :

1. Data yang digunakan dalam penelitian ini adalah data mitra binaan PKBL PT. Pusri dari tahun 2010 sampai dengan tahun 2017.
2. Penerapan teknik *data mining* dilakukan dengan menggunakan algoritma *Classification Based on Association (CBA)*.
3. Sistem yang akan dikembangkan adalah sebuah sistem klasifikasi risiko kredit calon mitra binaan PKBL PT. Pusri berbasis *web*.

DAFTAR PUSTAKA

- Addae-korankye, A. (2014). Causes and Control of Loan Default / Delinquency in Microfinance Institutions in Ghana. *American International Journal of Contemporary Research*, 4(12), 36–45.
- Amseke;, R., & Winarko, E. (2014). Aplikasi Algoritma CBA untuk Klasifikasi Resiko Pemberian Kredit, 8(2), 121–132.
- Bassil, Y. (2012). A Simulation Model for the Waterfall Software Development Life Cycle. *International Journal of Engineering & Technology*, 2(5), 2049–3444. <https://doi.org/10.15680/ijirce.2015.0305013>
- Bhasin, H., Khanna, E., & Sudha. (2014). Black Box Testing based on Requirement Analysis and Design Specifications. *International Journal of Computer Applications*, 87(1), 36. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.429.698&rep=rep1&type=pdf>
- Coleman, J. P. H. (2013). Data Flow Sequences: A Revision of Data Flow Diagrams for Modelling Applications using XML. *International Journal of Advanced Computer Science and Applications(IJACSA)*, 4(5), 28–31. Retrieved from <http://ijacsa.thesai.org/>
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. Morgan Kaufmann (Second). San Francisco. <https://doi.org/10.1007/978-3-642-19721-5>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Morgan Kaufmann (Third). Waltham. <https://doi.org/10.1016/B978-0-12-381479-1.00001-0>

- Heiderich, M. (2013). *Web Application Obfuscation. Journal of Chemical Information and Modeling* (Vol. 53).
<https://doi.org/10.1017/CBO9781107415324.004>
- Holmes, D. T. (2015). Cp-R, an interface the R programming language for clinical laboratory method comparisons. *Clinical Biochemistry*, 48(3), 192–195.
<https://doi.org/10.1016/j.clinbiochem.2014.10.015>
- Ibrahim, R., & Yen, S. Y. (2011). A Formal Model for Data Flow Diagram Rules. *Journal of System and Software*, 1(2), 60–69. Retrieved from scientific-journals.org/archive/vol1no2/vol1no2_3.pdf
- J. Zaki, M., & Meira, W. (2014). *Data Mining and Analysis: fundamental concepts and algorithms. Data Mining and Analysis.*
- Jafar Hamid, A., & Ahmed, T. M. (2016). Developing Prediction Model of Loan Risk in Banks Using Data Mining. *Machine Learning and Applications: An International Journal*, 3(1), 1–9. <https://doi.org/10.5121/mlaij.2016.3101>
- Jaffer, U., Vaughan-Huxley, E., Standfield, N., & John, N. W. (2013). Medical mentoring via the evolving world wide web. *Journal of Surgical Education*, 70(1), 121–128. <https://doi.org/10.1016/j.jsurg.2012.06.024>
- Kargarfard, F., Sami, A., & Ebrahimie, E. (2015). Knowledge discovery and sequence-based prediction of pandemic influenza using an integrated classification and association rule mining (CBA) algorithm. *Journal of Biomedical Informatics*, 57, 181–188.
<https://doi.org/10.1016/j.jbi.2015.07.018>
- Nadali, A., Kakhky, E. N., & Nosratabadi, H. E. (2011). Evaluating the success level of data mining projects based on CRISP-DM methodology by a Fuzzy

expert system. *ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology*, 6, 161–165.
<https://doi.org/10.1109/ICECTECH.2011.5942073>

Nguyen, L. T. T., Vo, B., Hong, T.-P., & Thanh, H. C. (2012). Classification based on association rules: A lattice-based approach. *Expert Systems with Applications*, 39(13), 11357–11366.
<https://doi.org/10.1016/j.eswa.2012.03.036>

Piparo, D., Tejedor, E., Mato, P., Mascetti, L., Moscicki, J., & Lamanna, M. (2018). SWAN: A service for interactive analysis in the cloud. *Future Generation Computer Systems*, 78, 1071–1078.
<https://doi.org/10.1016/j.future.2016.11.035>

Stanton, J. (2013). *Introduction to Data Science* (3rd ed.). New York: Syracuse University, School of Information Studies. <https://doi.org/10.1007/978-3-319-20424-6>

Watt, A., & Eng, N. (2014). *Database Design*. Retrieved from <https://opentextbc.ca/dbdesign01/>