

MODEL PEMBELAJARAN MESIN DALAM MERAMAL TAHAP LITERASI KEWANGAN ISI RUMAH KELAS PENDAPATAN T20, M40 DAN B40

Azmahanim binti Zamri

Azuraliza Abu Bakar

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Di Malaysia, terdapat tiga kelas pendapatan iaitu B40, M40 dan T20. Pendapatan isi rumah dikelaskan berdasarkan tiga petunjuk utama: pendapatan isi rumah, negeri, dan strata. Literasi perancangan kewangan sangat penting dalam menguruskan kewangan bagi menjamin kesejahteraan hidup. Perancangan kewangan yang berkesan membolehkan pengurusan kewangan dilakukan dengan baik. Literasi kewangan di kalangan warganegara Malaysia masih mencabar, terutamanya dalam penganggaran rumah tangga, menguruskan hutang, kredit, dan pelaburan. Oleh kerana kekurangan taksonomi ini, kelas B40 dan sebahagian M40 menghadapi beban kewangan yang telah dilupakan oleh kerajaan. Dalam kajian ini, kami mencadangkan untuk mengembangkan pendekatan pembelajaran mesin untuk memodelkan ramalan celik kewangan di kalangan kelas pendapatan di Malaysia. Kajian ini akan mengadaptasi metodologi CRISP-DM, yang terdiri dari beberapa fasa: pemahaman masalah, pengumpulan data, pra-pemrosesan dan persiapan data, pengembangan model ramalan, penilaian model, dan penerapan model. Kami akan menggunakan data yang ditinjau pada tahun 2016 yang diambil dari Jabatan Perangkaan Malaysia. Fokus utama kerajaan adalah golongan berpendapatan rendah sehingga meninggalkan golongan berpendapatan sederhana dalam krisis. Model ini memanfaatkan orang untuk mengkaji faktor dan corak penting yang menyumbang kepada celik kewangan dan kaedah psikologi kepada golongan berpendapatan sederhana. Penemuan faktor-faktor yang menyumbang kepada literasi kewangan kelas berpendapatan sederhana di Malaysia ini dapat membantu pemerintah Malaysia menggunakan model ini dalam menetapkan inisiatif dan dasar baru untuk celik kewangan golongan kelas berpendapatan sederhana dan rendah.

1 PENGENALAN

Literasi perancangan kewangan sangat penting dalam menguruskan kewangan bagi menjamin kesejahteraan hidup. Perancangan kewangan yang berkesan membolehkan

pengurusan kewangan dilakukan dengan baik. Keyakinan dan keupayaan untuk menguruskan kewangan ialah ciri hidup yang sangat diperlukan untuk mencapainya mengekalkan kualiti kehidupan yang baik. Mereka yang celik kewangan dapat membuat keputusan kewangan secara bijak pada setiap peringkat kehidupan ke arah peningkatan taraf hidup masing-masing.

Keupayaan kewangan membolehkan isi rumah berinteraksi secara bertanggungjawab dengan penyedia perkhidmatan kewangan, seterusnya menyumbang kepada daya tahan kewangan mereka. Kurangnya literasi perancangan kewangan menyebabkan berlakunya kegagalan menguruskan sumber kewangan yang boleh mengakibatkan berlakunya keruntuhan rumah tangga, masalah kesihatan seperti gangguan psikologi dan gangguan emosi. Rakyat Malaysia kini dibahagikan kepada tiga golongan, iaitu T20, M40, dan B40. Klasifikasi Pendapatan di Malaysia ini merupakan salah satu petunjuk pertumbuhan ekonomi, dan ternyata tahap setiap kumpulan pendapatan ini telah meningkat sejak beberapa tahun kebelakangan ini.

2 PENYATAAN MASALAH

Rakyat Malaysia memiliki tahap keyakinan yang rendah terhadap pengetahuan kewangan mereka sendiri yang boleh menimbulkan masalah perbelanjaan isi rumah mereka. 1 daripada 3 rakyat Malaysia mempunyai kurang ilmu pengetahuan yang diperlukan mengenai pengurusan kewangan. Antara kesan yang boleh dihadapi oleh isi rumah yang berpendapatan rendah mahu pun yang tinggi jika mereka mempunyai kurang ilmu kewangan, ianya mampu mendatangkan masalah kewangan yang kritikal pada masa hadapan.

Seharusnya lebih banyak institusi dan agensi lain perlu terbabit bagi menyampaikan mesej pendidikan kewangan dengan cara yang lebih meluas ke setiap lapisan masyarakat sehingga kadar literasi kewangan rakyat ataupun masyarakat dapat dipertingkatkan tetapi pada masa sekarang masih tidak banyak agensi yang berbuat sedemikian. Oleh yang demikian, berdasarkan pernyataan masalah

berhubung isutersebut, maka kaji selidik ini dilakukan untuk memperolehi maklumat mengenai literasi kewangan di kalangan kumpulan pendapatan T20, M40 dan B40. Kaedah yang lebih efektif perlu diaplikasikan bagi mengekstrak data mentah pendapatan isi rumah dalam setiap kumpulan tersebut dengan lebih tepat.

3 OBJEKTIF KAJIAN

Projek ini bertujuan Mengenal pasti faktor yang mempengaruhi tahap literasi kewangan golongankelas berpendapatan B40, M40 dan T20 menggunakan kaedah pemilihan fitur. Selain itu, mengkaji hubungan antara pendapatan dan corak perbelanjaan dengan literasi kewangan dalam kalangan rakyat Malaysia.

Kertas ini juga akan membincangkan pembangunan model ramalan dalam menentukan tahap literasi kewangan golongan kelas berpendapatan B40, M40 dan T20 menggunakan kaedah pembelajaran mesin.

4 METOD KAJIAN

Penggunaan model pembangunan yang sesuai penting untuk memastikan perjalanan projek berjalan dengan lancar dan menjamin hasil kerja yang berkualiti. Kajian ini akan mengadaptasi metodologi CRISP-DM (*cross-industry process for data mining*) bagi membangunkan model ramalan ini. CRISP-DM merupakan satustandard yang telah dikembangkan pada tahun 1996 yang ditujukan untuk melakukan proses analisis sebagai strategi pemecahan masalah dari bisnes atau unit penelitian. Justeru, projek yang dihasilkan akan menjadi lebih efisien kerana segala keperluanbaru dapat ditambah dengan mudah semasa fasa reka bentuk.

4.1 Fasa Pemahaman Bisnes

Fasa ini melibatkan proses memahami keperluan projek ini dari perspektif perniagaan, dan kemudian menukar pengetahuan ini ke dalam definisi masalah perlombongan data dan melakukan perancangan awal bagi mencapai objektif. Bagi

projek ini, masalah yang dihadapi ialah masalah literasi kewangan isi rumah dalam golongan kelas berpendapatan yang berbeza yang ada di Malaysia iaitu B40, M40 dan T20. Rakyat Malaysia memiliki tahap keyakinan yang rendah terhadap pengetahuan kewangan mereka sendiri yang boleh menimbulkan masalah perbelanjaan isi rumah mereka. Perkara ini akan mendatangkan masalah kewangan yang kritikal pada masa hadapan dan kerajaan akan memandang enteng sahaja masalah ini. Permasalahan ini menjadi indikator kepada kajian ini sebuah matlamat yang perlu diselesaikan. Sorotankajian dan kajian kesusasteraan yang telah dilakukan dapat memudahkan proses membangunkan projek ini. Selain itu, saiz data juga memainkan peranan yang besar di mana semakin besar saiz data, potensi menghasilkan keputusan yang baik semakin tinggi.

4.2 Fasa Pemahaman Data

Fasa pemahaman data bermula dengan pengumpulan data dan seterusnya memahami data, untuk mengenal pasti masalah data. Fasa pemahaman data adalah fasa ke dua dalam metodologi pembangunan model bagi kajian ini. Sumber data yang digunakan untuk kajian tahap literasi kewangan isi rumah kelas pendapatan yang berbeza di Malaysia ialah kaji selidik pendapatan dan perbelanjaan isi rumah 2016 yang diperoleh daripada Jabatan Statistik Malaysia. Sejumlah 14.5 juta data kaji selidik pendapatan dan perbelanjaan isi rumah pada tahun 2016 diperoleh. Pengumpulan data mentah perbelanjaan isi rumah dikutip bagi memahami situasi yang dialami dan membiasakan diri dengan data yang diperoleh.

4.3 Fasa Penyediaan Data

Fasa ini melibatkan proses penggabungan data dan pembersihan data. Fasa ini dijalankan bagi memudahkan fasa pembangunan model dijalankan bagi menghasilkan output yang tepat. Microsoft Excel dan Weka telah digunakan bagi melakukan pra- pemrosesan data. Fasa ini sangat berguna bagi memperoleh pengetahuan melalui langkah penyediaan dan penapisan data yang akan dipilih untuk digunakan dalam kajian ini. Antara proses yang dilalui ialah pembersihan data, pendiskretan data dan transformasi data.

Setelah melakukan pra pemrosesan ke atas data tersebut, tiga versi set data telah dikeluarkan di mana set data yang pertama ialah semua atribut yang telah melalui pra pemrosesan dipilih, set data yang ke dua ialah atribut dipilih mengikut tesis yang telah dirujuk iaitu tesis daripada Mohd Shahrum Nizam yang telah diberi oleh Dr Azuraliza sebagai bahan rujukan dan yang set data yang ketiga ialah pengurangan atribut yang di buat berdasarkan set data pertama dan ke dua.

4.4 Pemodelan

Bagi pelaksanaan kajian ini, data yang ada telah dipecahkan kepada tiga bentuk data yang berbeza. Setiap data set ini akan melalui proses pemodelan dan didapatkan pemboleh ubah yang terbaik bagi beberapa model iaitu pokok keputusan, hutan rawak, naif Bayes dan *Multilayer Perceptron*. Perisian yang digunakan untuk membangunkan model ialah Jupyter Notebook dan Google Colab.

Langkah-langkah yang diambil untuk membangunkan model setiap set data adalah seperti berikut:

1. Import perpustakaan yang akan digunakan untuk pembangunan model.
2. Memasukkan data yang digunakan dalam bentuk csv ke dalam perpustakaan pandas.
3. Memecahkan data kepada dua bentuk data dilakukan iaitu data untuk latihan dan data untuk peramalan bagi mengukur tahap ketepatan model yang dibangunkan. Nisbah yang diguna pakai untuk kajian ini adalah tujuh puluh peratus untuk latihandan tiga puluh peratus untuk ujian.
4. Algoritma model yang akan digunakan adalah pokok keputusan, hutan rawak, naif Bayes dan *Multilayer Perceptron*.

5 HASIL KAJIAN

Bahagian ini membincangkan hasil daripada proses pembangunan model ramalan tahap literasi kewangan isi rumah kelas pendapatan B40, M40 dan T20 di Malaysia.

Penerangan yang mendalam tentang pemodelan tahap literasi isi rumah diperihal. Dalam projek ini, model algoritma yang digunakan adalah Pokok Keputusan (*Decision Tree*), Hutan Rawak (*Random Forest*), naif Bayes (*Naïve Bayes*) dan *Multilayer Perceptron*. Pemodelan di buat ke atas tiga set data yang telah dikeluarkan dalam fasa pra pemrosesan.

5.1 Penilaian analisa kajian

Set data pertama

Model	Ketepatan	Purata
DT	0.60	0.65
RF	0.65	0.70
NB	0.65	0.71
MLP	0.66	0.71

Jadual 1 Gabungan semua model set 1

Berdasarkan jadual di atas, dapat diperhatikan bahawa MLP mempunyai ketepatan yang paling tinggi iaitu 0.66. Akan tetapi secara purata bagi pengukuran semua model ini mendapati bahawa nilai purata paling tinggi diperolehi ialah model NB iaitu 0.71.

Set data kedua

Model	Ketepatan	Purata
DT	0.62	0.67
RF	0.64	0.69
NB	0.65	0.71
MLP	0.68	0.73

Jadual 2 Gabungan semua model set 2

Daripada keputusan yang diperoleh dalam data set ini, model MLP mempunyai ketepatan yang tinggi berbanding dengan model yang lain. Secara purata bagi pengukuran semua model ini mendapati bahawa nilai purata paling tinggi diperolehi ialah juga model MLP iaitu 0.73.

Set data ketiga

Model	Ketepatan	Purata
DT	0.64	0.69
RF	0.65	0.71
NB	0.66	0.72
MLP	0.66	0.71

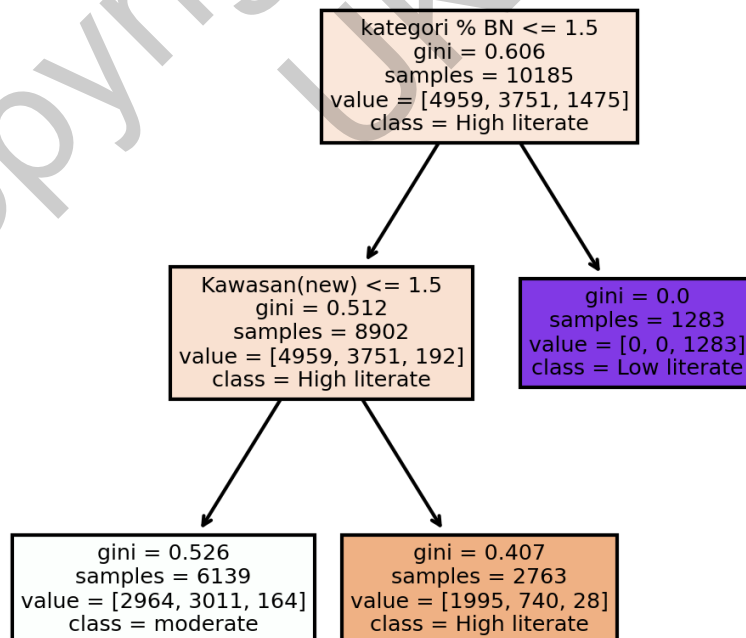
Jadual 3 Gabungan semua data set 3

Daripada keputusan yang diperoleh dalam data set ini, model NB dan MLP mempunyai ketepatan yang tinggi berbanding dengan model yang lain. Akan tetapi secara purata bagi pengukuran semua model ini mendapati bahawa nilai purata paling tinggi diperolehi ialah model NB iaitu 0.72.

5.2 Analisis petua pokok keputusan

Kaedah Algoritma DT bagi pembelajaran mesin untuk tujuan perlombongan data dapat mengembangkan grafik klasifikasi dan meramal kelas sasaran berdasarkan set data latihan input. Bagi kajian ini, pokok keputusan biasa atau DT telah digunakan sebagai algoritma perlombongan data.

Kaedah ini telah di buat ke atas set data yang terakhir iaitu set atribut gabungan daripada set 1 dan set 2. Rajah di bawah merupakan petua pokok keputusan yang dapat dihasilkan.



Rajah 1 Petua DT

Dari gambar rajah berkenaan berikut adalah jadual petua yang dapat dijana. Petua yang pertama ialah jika kategori peratus belanja bukan keperluan bersamaan dengan kategori 2 dan ke bawah dan kawasan yang didiami ialah di Semenanjung/Sabah, jadi taraf literasi isi rumah tersebut adalah sederhana. Petua yang kedua pula ialah jika kategori peratus belanja bukan keperluan bersamaan dengan kategori 2 dan ke bawah dan kawasan yang didiami ialah di Sarawak, jadi taraf literasi isi rumah tersebut adalah tinggi.

Berdasarkan petua DT di atas, dapat diperhatikan bahawa antara ciri-ciri yang menjadi sebab kepada tahap literasi kewangan sesuatu isi rumah ialah peratus perbelanjaan bukan keperluan dan lokasi isi rumah sama ada di semenanjung atau Sabah dan Sarawak.

6 KESIMPULAN

Pembangunan model pembelajaran mesin dalam meramal tahap literasi rakyat Malaysia dapat digunakan untuk menentukan tahap literasi sesuatu kategori kelas pendapatan iaitu B40, M40 dan T20. Dengan maklumat yang telah dihasilkan dalam kajian ini juga diharapkan agar dapat memberikan gambaran atau maklumat penting untuk menjalankan kajian lanjut. Kerajaan juga dapat memainkan peranan untuk membantu golongan M40 dan B40 yang mempunyai bebanan kewangan. Penyaluran bantuan dapat dilakukan secara lebih bersasar dikalangan rakyat Negara Malaysia.

Keseluruhannya, kajian ini telah berjaya mencapai kesemua objektif kajian seperti yang dinyatakan dalam Bab I. Kajian taraf literasi dalam kalangan rakyat Malaysia menggunakan pembelajaran mesin mempunyai maklumat dalam pelbagai aspek. Kajian ini merupakan sebuah kajian yang berasaskan algoritma pembelajaran mesin berasaskan beberapa model yang digunakan. Kelebihan pencapaian ini dapat menggalakkan lebih pengkaji menggunakan kajian ini bagi mendapatkan maklumat yang lebih spesifik dalam meneruskan kajian lanjut.

7 RUJUKAN

Atkinson, A., Lefevre, A. F., Chapman, M., Nassr, I. K., Robano, V., Wehinger, G., & Blundell-Wignall, A. 2012. OECD Working Papers on Finance, Insurance and Private Pensions.

Chamhuri S., Alias N. A., Ismail K. 2019. Kumpulan Rumah Berpendapatan 40 Peratus Terendah (B40) Di Malaysia: Mengenal Pasti Trend, Ciri, Isu dan Cabaran.

Cordero, J.M., M. Gil-Izquierdo, and F. Pedraja-Chaparroa 2019. Financial education and student financial literacy: A cross-country analysis using PISA 2012 data.

Hazrul Shahiri, Mohd Aqmin Abdul Wahab, Mustazar Mansur, Mohd Azlan Shah Zaidi. The rising cost of living in Malaysia: A slow hh income growth or increasing standard of living.

Nguyen, K. N. 2013. High School Seniors Financial Knowledge: the Impact of Financial Literacy Classes and Developmental Assets Thesis, University of Arkansas.