

Pengelasan Sebutan Huruf Hijaiyah menggunakan Teknik Pembelajaran Mesin

Nur Afini Natrah Mohd Ashril^a
p100933@siswa.ukm.edu.my
Fakulti Teknologi dan Sains Maklumat,
Universiti Kebangsaan Malaysia, Malaysia

Afzan Adam
afzan@ukm.edu.my
Fakulti Teknologi dan Sains Maklumat,
Universiti Kebangsaan Malaysia, Malaysia

Dahlila Putri Dahnil^b
dahlilaputri@ukm.edu.my
Fakulti Teknologi dan Sains Maklumat,
Universiti Kebangsaan Malaysia, Malaysia

ABSTRAK

Fitur *Mel-frequency cepstral coefficients* (MFCC) dan teknik pengelasan berdasarkan pembelajaran mesin sering digunakan dalam mengelaskan sebutan huruf-huruf hijaiyah. Walaupun begitu, berdasarkan kajian-kajian lepas, prestasi ketepatan pengelasan sebutan huruf hijaiyah masih lagi rendah walaupun dengan penggunaan algoritma pembelajaran mesin dan fitur MFCC. Oleh itu, kajian khas untuk menganalisis fitur dan teknik pembelajaran mesin yang sesuai akan dibincangkan dalam kertas kajian ini. Selain itu, bilangan huruf hijaiyah juga ditingkatkan kepada 30 huruf mengikut resam uthmani. Kajian ini mahu membuktikan bahawa penggunaan fitur dan teknik pengelasan yang sesuai mampu mengelaskan sebutan huruf hijaiyah dan memberikan prestasi ketepatan yang tinggi walaupun dengan jumlah huruf yang banyak. Kajian ini dilakukan berdasarkan kepada enam fasa utama dalam metodologi kajian ini termasuklah pemprosesan isyarat, penyarian fitur, pemprosesan dan pemilihan fitur, pengelasan dan akhir sekali pengujian, penilaian dan analisis. Kadar persampelan yang digunakan bagi kesemua modul pemprosesan isyarat pertuturan dalam kajian ini adalah 44.1 kHz. Dapatan kajian menunjukkan fitur MFCC merupakan fitur paling sesuai bagi mengelaskan sebutan huruf hijaiyah berbanding fitur-fitur lain yang telah diekstrak berdasarkan kepada 'rank' dalam hasil pemilihan fitur. Perbandingan prestasi ketepatan menunjukkan teknik pengelasan *Random Forest* (RF) mencapai ketepatan yang tinggi dengan menggunakan fitur MFCC, iaitu purata sebanyak 97~99% bagi setiap huruf hijaiyah berbanding teknik pengelasan lain yang telah diuji dalam kajian ini. Kesimpulannya, penggunaan fitur MFCC dan teknik pengelasan RF mampu memberikan prestasi ketepatan pengelasan sebutan huruf hijaiyah yang tinggi sekaligus meningkatkan prestasi pengelasan sebutan huruf hijaiyah kajian lepas, sehingga 98.29% secara purata untuk 30 huruf.

Kata kunci: Sebutan huruf hijaiyah; Pengelasan pertuturan; MFCC; Pembelajaran mesin; Pengecaman pertuturan

^a Penulis utama

^b Penulis koresponden

Classification of Hijaiyah Letters Pronunciation using Machine Learning Techniques

ABSTRACT

Mel-frequency cepstral coefficients (MFCC) features and classification techniques based on machine learning are often used in classifying hijaiyah letter pronunciations, however, the classification accuracy performance of hijaiyah letter pronunciations is still low even with the use of machine learning algorithms and MFCC features. Therefore, this study to analyze the features and relevant machine learning techniques will be presented in this study paper. In addition, the number of hijaiyah letters was also increased to 30 letters following the Uthmani resm. This research aims to prove that the suitable feature and relevant classification technique allows for precise classification of the pronunciation of each letter even with large amounts of letters. This research is conducted based on the six main stages in research methodologies which includes signal processing, feature searching, processing and feature selection, classification and lastly, testing, evaluation and analysis. The sampling rate used for all speech signal processing modules in this study is 44.1 kHz. The findings of the study show that the MFCC feature is the most suitable feature to classify the pronunciation of hijaiyah letters compared to other features that have been extracted based on the rank in the feature selection results. Comparison of accuracy performance shows that Random Forest (RF) classification technique achieves high accuracy by using MFCC feature, which is an average of 97 ~ 99% for each hijaiyah letter compared to other classification techniques that have been tested in this study. In conclusion, the use of MFCC features and RF classification techniques are able to provide a high performance of hijaiyah pronunciation classification accuracy, which is 98.29% on average even with the use of 30 letters.

Keywords: Hijaiyah letters pronunciation; Speech classification; MFCC; Machine learning; Speech recognition

PENGENALAN

Huruf Arab atau huruf hijaiyah ialah huruf yang membentuk perkataan dalam al-Quran yang terdiri daripada 30 huruf daripada huruf ^ا hingga ^ي. Penguasaan terhadap sebutan huruf hijaiyah boleh ditakrifkan sebagai kecekapan seseorang menyebut huruf tersebut mengikut makhraj huruf masing-masing. Makhraj adalah perkataan bahasa Arab bagi tempat keluar sesuatu sebutan huruf berdasarkan (Nurul Wahidah Arshad et al., 2013). Makhraj merujuk kepada kedudukan organ pertuturan yang menghasilkan huruf supaya dapat dibezakan dengan huruf yang lain, menurut (Nurul Wahidah Arshad et al., 2011). Dalam membaca al-Quran, pembaca wajib membaca dengan makhraj yang betul. Hal ini kerana walaupun sedikit perbezaan pada bunyi perkataan Arab, ianya boleh membawa kepada makna yang berbeza.

Pengecaman pertuturan menjadi salah satu tarikan penting telefon pintar hari ini kerana mempunyai fungsi sebagai pembantu maya seperti *Alexa*, *Cortana* dan *Siri*. Menurut (Ben & Karolina, 2021), dalam *Apple iPhone*, pengecaman pertuturan menguasai papan kekunci dan memudahkan pengguna untuk menaip tanpa menggunakan papan kekunci. Tambahan lagi, pengecaman pertuturan juga boleh didapati dalam aplikasi pemprosesan perkataan seperti *Microsoft Word*, di mana pengguna boleh menukar perkataan kepada teks.

Kewujudan pembelajaran bermakna (*deep learning*) sebagai lanjutan kepada teknologi pembelajaran mesin telah membuatkan pengecaman pertuturan mampu untuk mengenali pertuturan dengan lebih tepat (Hiriyannaiah et al., 2020). Menurut (Anusuya & Katti, 2009), pengecaman pertuturan merupakan proses menukar isyarat pertuturan kepada susunan perkataan menggunakan algoritma yang diimplementasikan sebagai program komputer. Selain itu, (Moawad, 2012) menerangkan bahawa proses menukar isyarat akustik menjadi satu set perkataan juga merupakan definisi kepada pengecaman pertuturan. Seiring dengan perkembangan zaman, teknologi juga turut berkembang. Proses mempelajari al-Quran menjadi semakin mudah (Siti Nor Azimah Sabaruddin & Tengku Intan Zarina Tengku Puji, 2013). Kemajuan teknologi pertuturan hari ini sedikit sebanyak menjadi alat bantu proses pembelajaran pengecaman huruf abjad, pengecaman nombor dan juga pengecaman huruf arab yang dikenali sebagai huruf hijaiyah. Teknik pengecaman pertuturan sangat diperlukan bagi pengecaman pertuturan huruf hijaiyah bagi proses pembelajaran. Ini kerana huruf hijaiyah menjadi asas bagi membentuk perkataan-perkataan dalam al-Quran.

Huruf hijaiyah ialah huruf-huruf yang membentuk kata-kata dalam al-Quran yang terdiri daripada 28 huruf secara asasnya dan melambangkan bunyi konsonan. Manakala bunyi vokal pula diwakili oleh harakat/tanda baca. Harakat atau tanda baca digunakan untuk menjelaskan sebutan huruf-huruf tersebut (Untari et al., 2017). Penguasaan sebutan huruf hijaiyah boleh didefinisikan sebagai kecekapan seseorang menyebut huruf hijaiyah mengikut makhrajnya. Pada peringkat permulaan pembelajaran al-Quran, penguasaan sebutan makhraj huruf hijaiyah perlu ditekankan, kerana tanpa penguasaan yang baik, akan membawa kepada perubahan makna yang berlainan. Penekanan kepada penguasaan sebutan huruf hijaiyah perlu diberi penekanan pada peringkat awal pembelajaran bahasa Arab lagi (Rahimi et al., 2010). Walau bagaimanapun, terdapat beberapa pandangan alim ulama' berkenaan jumlah huruf hijaiyah yang boleh digunapakai kerana terdapat sedikit variasi dari segi tulisan dan sebutan setiap huruf (Teena, 2019). Ada yang berpendapat huruf hijaiyah terdiri daripada 28 dan 29 huruf. Jika berdasarkan al-Quran Muqaddam Resam Uthmani yang digunakan secara meluas di Malaysia sebagai bahan rujukan bagi pembelajaran huruf hijaiyah, terdapat 30 huruf hijaiyah secara keseluruhannya seperti yang dinyatakan oleh (Nurul Wahidah Arshad et al., 2011).

al-Quran merupakan kitab suci bagi umat Islam dan ditulis dalam Bahasa Arab. al-Quran merupakan mukjizat terakhir yang telah diturunkan kepada Nabi Muhammad SAW bagi memimpin umat Islam (al-Qattan, 1973). Ayat-ayat al-Quran disusun menggunakan huruf hijaiyah, maka sesiapa yang ingin membaca al-Quran haruslah mampu mengenal dan membaca huruf hijaiyah. Proses pembelajarannya secara konvensional memerlukan bimbingan secara bersemuka daripada guru untuk menunjukkan cara sebutan setiap huruf (Putra et al., 2012).

Pengecaman pertuturan merupakan salah satu bidang kajian teras dalam pemprosesan pertuturan dan juga dikenali sebagai *Automatic Speech Recognition (ASR)*. Ia adalah proses menukar isyarat pertuturan kepada urutan perkataan (iaitu, perkataan yang dituturkan kepada teks) melalui algoritma yang dilaksanakan sebagai program komputer (Xuedong et al., 2001).

Berdasarkan (Nilsson, 2013), pengecaman pertuturan ialah proses menukar aliran akustik input pertuturan, kepada perwakilan teks bagi komponen-komponen perkataan. Pengecaman pertuturan menggunakan bidang penyelidikan sains komputer, linguistik dan kejuruteraan komputer dengan meluas. Kebanyakan peranti moden dan program berfokuskan teks hari ini mempunyai fungsi pengecaman pertuturan di dalamnya untuk membolehkan pengguna menggunakan peranti bebas tangan (Ben & Karolina, 2021).

Berdasarkan kajian-kajian lepas, terdapat perbezaan dalam komponen rangka kerja pengecaman pertuturan yang dicadangkan dan digunakan dalam pelbagai bidang kajian. Walau bagaimanapun, kesemua kajian ini mempunyai motif yang sama tetapi cara yang sedikit berbeza bagi mencapainya bergantung kepada keperluan kajian yang dijalankan. (Chavan & Gawande 2015; Gevaert et al. 2010) bersetuju mengatakan bahawa rangka kerja asas bagi sistem pengecaman pertuturan bermula dengan input isyarat pertuturan samada menggunakan pangkalan data sedia ada atau membina pangkalan data baharu dengan mengumpulkan sampel suara. Kemudian, pemprosesan isyarat pertuturan diikuti dengan penyarian fitur dan akhir sekali pengelasan isyarat pertuturan tersebut.

Pengelasan menurut (Sonia & David Peter, 2013) bermaksud mengkategorikan data dan membentuk kumpulan berdasarkan persamaan yang dimiliki. Di samping itu, pengelasan juga merupakan salah satu teknik dalam pembelajaran mesin yang sering digunakan untuk meramal ahli bagi sesuatu kumpulan. Memandangkan terdapat kelas yang berbeza dalam setiap tugas pengecaman pertuturan, algoritma pengelasan binari atau berbilang kelas diperlukan. Dalam sesetengah kes, teknik pembelajaran mesin boleh dianggap sebagai salah satu jalan terbaik, dalam tugas mengecam corak pertuturan dengan tepat (Zitouni, 2014).

Salah satu isu dalam kajian pengelasan sebutan huruf-huruf hijaiyah yang lepas ialah fitur MFCC sering digunakan tetapi masih belum berjaya membuktikan bahawa ia mampu mewakili perbezaan antara huruf hijaiyah yang banyak kerana prestasi ketepatan yang terbaik hanyalah dengan tiga huruf sahaja berdasarkan kajian (Elvira, 2018). Walaupun fitur MFCC telah digunakan dan pelbagai teknik pengelasan telah diuji, belum mampu mengekalkan prestasi ketepatan yang baik apabila diuji dengan jumlah kelas (huruf) yang banyak. Ringkasnya, terdapat keperluan bagi menjalankan kajian bagi meningkatkan prestasi ketepatan pengelasan dalam mengelaskan sebutan 30 huruf hijaiyah dengan menggunakan kaedah penyarian dan pemilihan fitur serta dengan penggunaan teknik pembelajaran mesin.

(Fadhilah Rosdi et al., 2019) menerangkan bahawa MFCC merupakan fitur pertuturan yang mampu menangkap ciri penting bagi isyarat audio. MFCC telah digunakan secara meluas dalam bidang pengecaman pertuturan. Terdapat beberapa penyelidik yang telah menggunakan fitur MFCC bagi mengelaskan sebutan huruf-huruf hijaiyah (Adiwijaya et al., 2017; Elvira, 2018; Wisnu Adi et al., 2018).

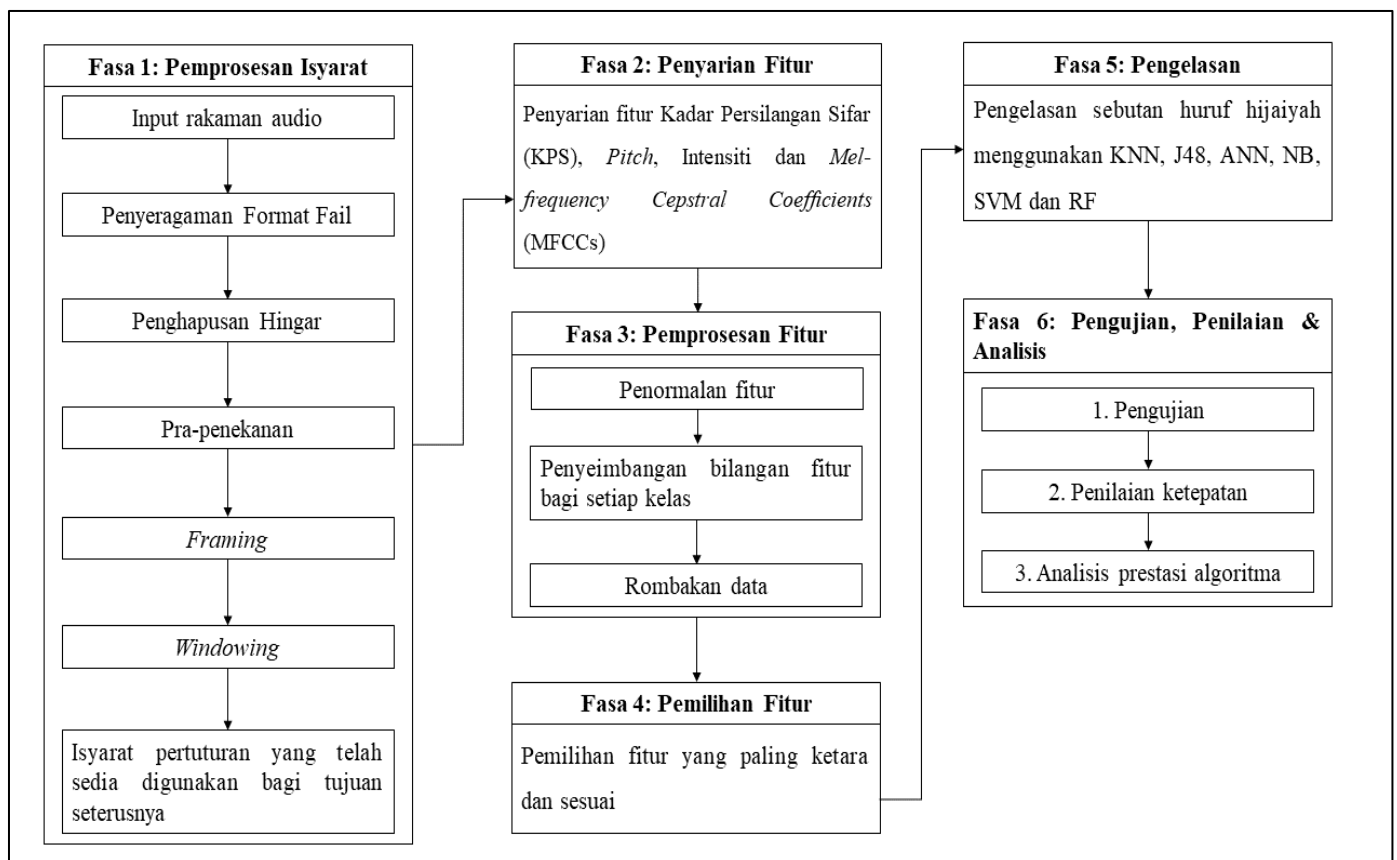
Hal ini kerana, fitur MFCC dipercayai meniru sistem pendengaran manusia dan dijangka menjadi fitur yang terbaik dan paling sesuai untuk diekstrak daripada sebutan huruf-huruf hijaiyah (Zbynek & Josef, 1999). Penyarian fitur merupakan proses mengenalpasti fitur penting dalam isyarat dan yang boleh digunakan dengan mudah oleh algoritma pembelajaran mesin berbanding dengan penggunaan data mentah (Anon, 2020). MFCC digunakan kerana ia mampu memberikan prestasi ketepatan yang baik dan juga kerumitan pengkomputeran yang kurang berbanding dengan kepada kaedah penyarian fitur yang lain, menurut (Rajeev & Rajesh Kumar, 2016).

Menurut (Xuedong & Li, 2009), ketepatan pengecaman pertuturan automatik kekal sebagai salah satu cabaran penyelidikan yang paling penting selepas bertahun-tahun penyelidikan dan pembangunan. Oleh itu, kajian ini dijalankan bertujuan untuk mengenalpasti fitur yang sesuai dan yang diperlukan antara fitur *pitch*, intensiti, kadar persilangan sifar dan MFCC bagi mengelaskan sebutan 30 huruf hijaiyah. Selain itu, kajian ini juga dibuat bagi mengenal pasti teknik pengelasan yang sesuai dan yang terbaik bagi mengecam sebutan 30 huruf hijaiyah seterusnya membuat penilaian serta perbandingan prestasi ketepatan pengelasan.

Artikel ini disusun seperti berikut: bahagian pertama merupakan pengenalan kepada topik kajian, diikuti dengan metodologi dan dapatan kajian. Bahagian terakhir meringkaskan hasil daripada kajian sebagai kesimpulan.

METODE

Metodologi kajian memberi gambaran yang lebih jelas dan menyeluruh tentang sesuatu kaedah yang dijalankan. Seksyen metode akan membincangkan secara terperinci proses yang terlibat secara langsung dalam mengelaskan sebutan 30 huruf hijaiyah berdasarkan penyarian serta pemilihan fitur dan teknik pengelasan. Metode kajian ini terdiri daripada enam fasa utama dimulai dengan fasa pemrosesan isyarat, penyarian fitur, pemrosesan fitur, fasa pengelasan dan akhir sekali pengujian, penilaian dan fasa analisis, seperti yang ditunjukkan dalam Rajah 1 di bawah.



RAJAH 1. Cadangan pembangunan proses penghasilan model pengecaman sebutan 30 huruf hijaiyah

FASA 1: PEMROSESAN ISYARAT

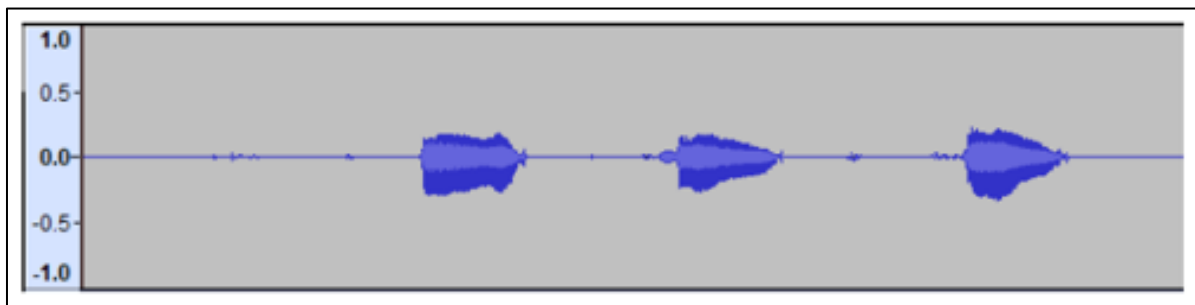
Berdasarkan (Al-Sabri et al., 2018), pra-pemrosesan merujuk kepada segala proses yang dilakukan kepada isyarat. Pra-pemrosesan merupakan suatu proses awal yang penting dalam menjalankan kajian berkaitan pembelajaran mesin. Pra-pemrosesan membantu dalam meningkatkan ketepatan membuat pengelasan, menurut (Ezuana Sukawai & Nazlia Omar,

2020). Langkah pertama ini melibatkan input rakaman audio, penyeragaman format fail, penghapusan hingar, pra-penekanan, *framing* dan *windowing*.

Dalam kajian ini, rakaman audio dikumpulkan dengan menggunakan Borang Google®. Borang tersebut dibina dan diedarkan kepada pelbagai pihak. Senarai 30 huruf hijaiyah yang perlu disebut dan cara sebutannya telah dimasukkan ke dalam borang tersebut sebagai rujukan responden. Responden memuatnaik satu fail rakaman suara yang mengandungi sebutan 30 huruf hijaiyah mengikut turutan. Data yang telah berjaya dikumpulkan melibatkan tiga orang lelaki dan 14 orang perempuan berusia lingkungan 18 sehingga 55 tahun, kesemuanya adalah warganegara Malaysia, seperti yang ditunjukkan dalam Rajah 4 di bawah.



RAJAH 2. Isyarat sebelum dinyah hingar

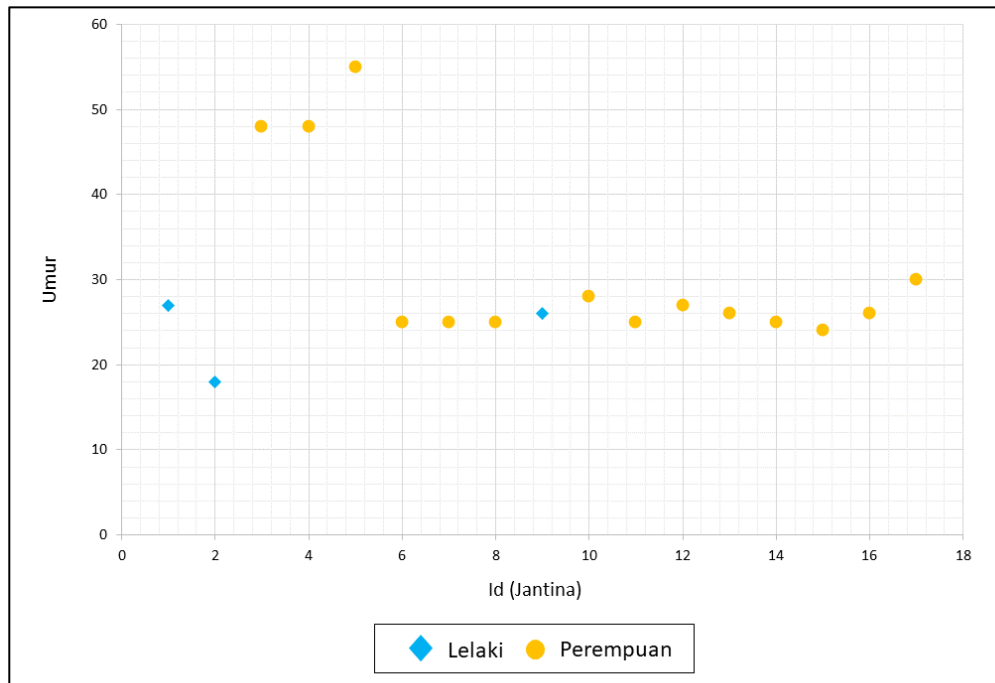


RAJAH 3. Isyarat selepas dinyah hingar

Langkah seterusnya dalam fasa pertama ini merupakan penyeragaman format fail. Fail rakaman suara yang dimuat naik oleh responden terdiri daripada pelbagai format seperti mp3, mp4a, amr dan 3gpp bagi memudahkan tugas responden. Penyeragaman format fail dilakukan dengan mengubah format asal sesuatu fail kepada format bentuk gelombang (juga dipanggil sebagai fail wav) sahaja. Menurut (Oksana, 2021), fail WAV adalah salah satu audio digital yang lebih popular format dan piawaian emas dalam sebarang rakaman studio dan mudah dikenali dan diproses oleh pelbagai perkakasan dan perisian.

Langkah berikutnya merupakan penghapusan hingar. Disebabkan data yang diperolehi dalam kajian ini dikumpul daripada pelbagai lokasi yang berbeza, setiap persekitaran lokasi mempunyai hingar yang tersendiri. Sesetengah sampel suara mengandungi hingar yang minimum, kerana direkod dalam bilik kedap bunyi tanpa kehadiran gema dan suara latar. Sesetengahnya lagi mungkin direkod dalam persekitaran yang mempunyai hingar seperti bunyi kipas, bunyi kenderaan lalu lalang mahupun bunyi haiwan seperti cengkerik yang tidak dapat dielakkan. Banyak kajian empirikal telah menunjukkan bahawa hingar dalam set data membawa kepada penurunan prestasi ketepatan klasifikasi. Masalah mengenal pasti dan mengendalikan

hingar telah menarik perhatian ramai sejak beberapa tahun yang lalu, berdasarkan (Shivani & Atul, 2019). Oleh itu, penghapusan hingar perlu dilakukan bagi mendapatkan audio yang berkualiti untuk menjamin sistem yang tepat dan boleh dipercayai. Penghapusan hingar telah dibuat menggunakan perisian Audacity®. Perbezaan sebelum dan selepas hingar dinyahkan ditunjukkan dalam Rajah 2 dan Rajah 3 di atas.



RAJAH 4. Profil responden kajian ini

Kemudian, pra-penekanan dijalankan ke atas kesemua fail rakaman suara bagi menambah baik kualiti isyarat ketika penghantaran gelombang bunyi. Dalam kajian ini, pra-penekanan dijalankan ke atas semua fail bagi mencapai keseragaman yang akan menghasilkan audio yang lebih berkualiti. Secara ringkasnya, penapis pra-penekanan sebenarnya meningkatkan komponen frekuensi tinggi sahaja, sambil membiarkan komponen frekuensi rendah dalam keadaan asalnya. Prosedur pra-penekanan dibuat menggunakan perisian Praat®, versi keenam (Weenink, 2018). Proses ini dijalankan menggunakan persamaan matematik 1 di bawah (Lindasalwa Muda et al., 2010).

$$y(n) = x(n) - ax(n-1)$$

Berdasarkan persamaan matematik 1 di atas, kebiasaannya, isyarat berfrekuensi rendah yang disampel pada kadar yang cukup tinggi, berpotensi untuk menghasilkan sampel berhampiran dengan nilai berangka yang serupa. Hal ini disebabkan oleh frekuensi rendah yang pada asasnya merujuk kepada variasi masa yang perlahan dan oleh itu nilai berangka isyarat frekuensi rendah cenderung berubah secara perlahan ataupun lancar dari sampel ke sampel. Dengan penolakan (-), bahagian sampel yang tidak berubah dan yang berhubung dengan sampel berhampiran dihapuskan dan yang tinggal ialah bahagian isyarat yang berubah dengan cepat (contoh: komponen frekuensi tinggi) (Nate, 2012).

Langkah terakhir bagi fasa pertama metode kajian ini ialah *framing* dan *windowing*. Langkah ini dilakukan kerana responden hanya memuat naik satu fail ke Borang Google® yang mengandungi kesemua 30 huruf hijaiyah. Bagi tujuan memproses satu-satu huruf untuk dikelaskan, audio tersebut perlu dipotong mengikut huruf dan dimasukkan ke dalam fail yang berasingan. Suara dipotong menjadi 30 *frame* berdasarkan 30 huruf hijaiyah. Proses pemotongan dijalankan menggunakan Audacity®. Perisian Audacity® melakukan proses pembingkai (*framing*) sendiri, berdasarkan parameter input daripada pengguna. Hasil akhir bagi fasa pertama ini akan digunakan dalam fasa seterusnya.

FASA 2: PENYARIAN FITUR

Doshi, 2018 menerangkan bahawa penyarian fitur merupakan fasa terpenting dalam menilai dan bagi mengenalpasti hubungan antara pelbagai objek. Fasa ini penting kerana data pendengaran yang diperolehi tidak dapat difahami secara langsung oleh model, oleh itu penyarian fitur digunakan untuk mengubahnya menjadi format yang boleh difahami. Terdapat empat fitur yang telah diekstrak bagi kajian ini termasuklah *Mel-frequency Cepstral Coefficients* (MFCC), Kadar Persilangan Sifar (KPS), intensiti dan *pitch*.

MFCC merupakan fitur yang kebiasaannya digunakan bagi pemprosesan pertuturan. Fitur MFCC juga digunakan secara meluas dalam kajian pengelasan sebutan huruf-huruf hijaiyah berdasarkan kajian-kajian lepas (Adiwijaya et al., 2017; Elvira, 2018; Wisnu Adi et al., 2018).

Setiap huruf menghasilkan 12 pekali MFCC. Purata setiap fitur MFCC tersebut digunakan kerana bilangan fitur yang diekstrak oleh perisian analisis pertuturan seperti Praat® kebiasaannya tidak sama panjang bagi setiap fail. Ini kerana fail input tidak mempunyai tempoh yang sama panjang. Oleh itu, purata pekali MFCC dijadikan input kepada model pembelajaran mesin. Fitur ini telah diekstrak menggunakan perisian Praat®.

Fitur seterusnya yang diekstrak dalam kajian ini ialah Kadar Persilangan Sifar (KPS). KPS merupakan ukuran bilangan dalam selang masa/bingkai masa tertentu apabila amplitud isyarat pertuturan melalui nilai sifar (Bachu et al., 2010). Dalam kajian ini, KPS diekstrak menggunakan perisian Jupyter Notebook® versi 6.1.4 dalam persekitaran Anaconda®.

Seterusnya, penyarian fitur keamatan. Berdasarkan kajian lepas (Mansour et al., 2011), fitur intensiti digunakan untuk mengelaskan jantina berdasarkan pertuturan kerana kemaatan suara lelaki dan perempuan yang amat jelas berbeza. Walaubagaimanapun, bagi kajian ini, fitur intensiti dipilih bagi melihat sejauh mana fitur ini mampu mengelaskan sebutan huruf-huruf hijaiyah.

Jumlah tenaga yang dihantar melalui ruang unit tertentu per unit masa dikenali sebagai intensiti gelombang bunyi (ThePhysicsClassroom, 2021). Perubahan tekanan dan amplitud adalah berkadar dengan intensiti (Schuster, 2013). Semakin tinggi intensiti, semakin kuat bunyi yang didengar dan semakin rendah intensiti, semakin perlahan bunyi yang didengar terang (Graham, 2014).

Fitur *pitch* merupakan fitur yang keempat diekstrak bagi kajian ini. *Pitch* ialah tempoh asas isyarat pertuturan. Ia merupakan perkaitan persepsi frekuensi asas. Ia melambangkan frekuensi getaran pita suara semasa pengeluaran bunyi. Salah satu sebab yang menjadikan fitur *pitch* menarik ialah fitur ini agak kebal terhadap modulasi frekuensi. Contohnya, apabila isyarat melalui penapisan laluan rendah. Jika isyarat dipindahkan melalui telefon, *pitch* isyarat tersebut akan kekal sama, hanya *pitch* anggaran yang akan terkesan. Ini berkemungkinan menjadikan fitur *pitch* lebih sesuai bagi tugas pengelasan (Andreas et al., 2006). Rakaman audio ini

dianalisis menggunakan perisian Praat[®], bagi memperolehi fitur intensiti (Kaseh Abu-bakar & Muhammad Faiz Abdullah, 2018) dan fitur *pitch*.

FASA 3: PEMROSESAN FITUR

Selepas kesemua fitur selesai diekstrak, ia perlu diproses bagi memastikan kualiti data terjamin dan seterusnya pengelasan yang dibuat mampu memberikan prestasi ketepatan yang baik. Data yang bersih akan mampu menambah baik produktiviti, menurut Tableau, 2021. (Mary, 2019) menambah, *garbage in, garbage out*. Jika baik data yang digunakan, baik jugalah hasilnya.

Penormalan fitur ialah proses menukar saiz satu atau lebih atribut. Penormalan fitur dibuat bagi memastikan data berada dalam julat yang sama. X_{max} dan X_{min} dalam persamaan matematik 2 di bawah adalah nilai maksimum dan minimum bagi fitur-fitur tersebut. Setiap atribut dalam data yang telah dinormalkan akan berada pada skala antara julat yang lebih kecil iaitu 0 hingga 1 atau -1 hingga 1.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Langkah seterusnya dalam pemprosesan fitur ialah penyeimbangan bilangan fitur bagi setiap kelas. Berdasarkan (Ezuana Sukawai, 2020), hal ini kerana ketidak seimbangan pengagihan data antara polariti boleh menjadi masalah bagi teknik pengelasan pembelajaran mesin. Selepas data dinormalkan, data yang tidak seimbang akan diseimbangkan. Data yang tidak seimbang akan mempengaruhi prestasi sesuatu teknik pengelasan. Prestasi sesuatu teknik pengelasan boleh menjadi berat sebelah akibat dipengaruhi oleh kelas majoriti. Oleh itu, masalah data yang tidak seimbang perlu diatasi terlebih dahulu sebelum diaplikasikan ke atas mana-mana teknik pengelasan. (Ye & Rick, 2017) mengatakan bahawa terdapat beberapa kaedah yang boleh digunakan bagi menyeimbangkan data termasuklah menggunakan metrik penilaian yang bersesuaian dan juga persampelan semula data yang hendak digunakan. Persampelan semula data melibatkan samada pengumpulan berkurang, pengumpulan berlebihan mahupun gabungan kedua teknik tersebut (Jason, 2019).

Kajian ini menggunakan Kaedah *Synthetic Minority Oversampling Technique* (SMOTE) bagi penyeimbangan data. SMOTE adalah teknik pengumpulan berlebihan bukan rawak di mana sampel sintetik dihasilkan khas bagi kelas minoriti (Swastik, 2020). (Cornellius, 2020) menerangkan bahawa pengumpulan berlebihan merujuk kepada menyeimbangkan jumlah data pada bahagian kelas minoriti agar sama jumlahnya dengan bahagian kelas majoriti. (Satyam, 2021; Cornellius, 2020) berpendapat SMOTE berfungsi menggunakan teknik KNN bagi membuat data sintetik. SMOTE terlebih dahulu memilih data secara rawak dari kelas minoriti, kemudian lipatan k terdekat dari data akan ditetapkan. Data sintetik kemudian dibuat antara data rawak dan jiran terdekat yang dipilih secara rawak. Prosedur diulang-diulang sehingga kelas minoriti mempunyai jumlah yang sama dengan kelas majoriti (Satyam, 2021).

Langkah terakhir bagi pemprosesan data fitur dalam kajian ini ialah rombak data. Data perlu dirombak agar tidak mengandungi lipatan yang berat sebelah ketika hendak digunakan dalam pengesanan silang lipatan k . Penapis jenis *Randomize* dipilih terlebih dahulu kemudian digunakan ke atas data. Perisian Weka[®] digunakan bagi menjalankan penyeimbangan, penormalan dan rombak data dalam kajian ini. Perisian atau alat analisis Weka[®] (Waikato Environment for Knowledge) merupakan perisian sumber terbuka perlombongan data yang

ditulis menggunakan Java dan boleh digunakan dalam pelbagai platform (Azhar Mohd Khairy et al., 2018).

FASA 4: PEMILIHAN FITUR

Pemilihan fitur yang relevan mempengaruhi peningkatan prestasi pengelasan, berdasarkan (Fadhilah Rosdi et al., 2019). Oleh itu, bagi kajian ini, pemilihan fitur dibuat sebagai salah satu langkah bagi menambah baik prestasi ketepatan pengelasan yang dijalankan supaya fitur yang digunakan betul-betul sesuai bagi mengelaskan sebutan 30 huruf hijaiyah. Langkah ini melibatkan pemilihan fitur untuk mendapatkan hasil terbaik daripada model pengelasan. Fitur yang bagus dan terbaik membantu untuk menunjukkan struktur asas maklumat. (Atul, 2020) menerangkan, dalam pembelajaran mesin, tujuan pemilihan fitur dibuat adalah bagi mengenalpasti set fitur yang terbaik bagi membangunkan model terbaik bagi sesuatu kajian. Ia juga boleh mengurangkan jumlah fitur dalam sesuatu model (Gabriel, 2019). (Fadhilah Rosdi et al., 2019) menambah, bilangan fitur yang banyak akan menambah masa pemprosesan pengelasan dan teknik pengelasan terdedah kepada *overfitting*. Rangka kerja umum pemilihan fitur adalah seperti *pseudocode* dalam Rajah 5 di bawah.

```
AttributeSelection attsel = new AttributeSelection();  
// bina dan tetapkan contoh AttributeSelection yang baru  
Ranker search = new Ranker();  
// pilih kaedah carian  
PrincipalComponents eval = new PrincipalComponents();  
// pilih kaedah penilaian  
attsel.setEvaluator(eval);  
// tetapkan kaedah penilaian  
attsel.setSearch(search);  
// tetapkan kaedah carian  
attsel.SelectAttributes(data);  
// tetapkan data yang hendak digunakan bagi pemilihan fitur
```

RAJAH 5. Pseudocode rangka kerja pemilihan fitur

Walau bagaimanapun, penggunaan fitur pertuturan yang tidak berkaitan pula akan menurunkan prestasi pembelajaran teknik pengelasan. Terdapat tiga jenis kaedah pemilihan fitur termasuklah kaedah *filter*, *wrapper* dan *embedded* (Dimitris, 2021).

Bagi kajian ini, pemilihan fitur telah dijalankan dengan menggunakan teknik *CorrelationAttributeEval* dengan kaedah carian *Ranker* dalam perisian Weka[®]. Menurut (Jason, 2016), perisian Weka[®] menyokong pemilihan fitur berdasarkan perkaitan menggunakan teknik *CorrelationAttributeEval* dengan kaedah carian *Ranker*. Tambahan lagi, teknik pemilihan fitur *CorrelationAttributeEval* menilai fitur dengan teknik carian *Ranker* bagi menunjukkan kepentingan setiap fitur berdasarkan kedudukan masing-masing.

Mengukur nilai fitur dengan menggunakan kaedah menaksir hubungan (*Pearson*) antara atribut dengan kelas. Perkaitan pekali (*Pearson*) boleh digunakan bagi menilai hubungan antara lebih daripada dua pemboleh ubah.

FASA 5: PENGELASAN

Dalam kajian ini, teknik pengelasan yang diaplikasikan adalah bertujuan untuk mengelaskan sebutan 30 huruf hijaiyah. Pengelasan dibuat berdasarkan fitur yang telah diekstrak dan dipilih. Tetapan parameter bagi pengelasan sebutan huruf-huruf hijaiyah dalam kajian ini adalah seperti dalam Jadual 1 di bawah.

JADUAL 1. Tetapan parameter teknik-teknik pengelasan yang digunakan dalam kajian ini

Teknik Pengelasan	Nilai Parameter
<i>K-Nearest Neighbour (KNN)</i>	<i>KNN=1</i> <i>batchSize= 100</i> <i>nearestNeighbourSearchAlgorithm= LinearNNSearch</i>
<i>J48 Trees</i>	<i>batchSize= 100</i> <i>numFolds= 3</i> <i>seed= 1</i>
<i>Artificial Neural Network (ANN)</i>	<i>batchSize= 100</i> <i>hiddenLayers= a</i> <i>seed= 0</i>
<i>Naïve Bayes (NB)</i>	<i>batchSize= 100</i>
<i>Support Vector Machine (SVM)</i>	<i>batchSize= 100</i> <i>kernel= Polynomial</i> <i>c= 1.0</i> <i>randomSeed= 1</i> <i>toleranceParameter= 0.001</i>
<i>Random Forest (RF)</i>	<i>batchSize= 100</i> <i>numIterations= 100</i> <i>seed= 1</i>

Pengelasan dijalankan dengan menggunakan kaedah lima kali lipatan silang ($k=5$). Pembahagian pengelasan silang dapat membantu mengurangkan peluang *overfitting* sesebuah model. Pengelasan silang berfungsi dengan cara merombak dan membahagikan secara rawak set data ke dalam kumpulan berjumlah k . Bagi setiap kumpulan, sebahagian set data akan digunakan sebagai set pengelasan manakala sebahagian lagi akan menjadi set latihan. Oleh itu, setiap data latihan akan menjadi data ujian dan juga sebaliknya. Pengelasan akan pelajari daripada set latihan dan akan diaplikasikan kepada set pengelasan yang belum pernah dilihat lagi.

Proses ini berulang bagi kesemua kumpulan dan ketepatan yang dihasilkan digabungkan dan purata ketepatan akan dipilih sebagai ketepatan akhir dengan cara merombak dan membahagikan secara rawak set data ke dalam kumpulan berjumlah k .

Terdapat dua kaedah eksperimen dijalankan dalam kajian ini bagi menguji pengaruh/keberkesanan fitur terhadap pengelasan. Eksperimen pertama menggunakan pengelasan fitur tunggal. Eksperimen kedua pula menggabungkan dua fitur.

Fitur MFCC, KPS, intensiti dan *pitch* yang telah diekstrak akan diuji secara berasingan ke atas enam teknik pengelasan berasaskan pembelajaran mesin yang telah dipilih iaitu KNN, J48, ANN, NB, SVM dan RF. Berdasarkan prestasi ketepatan yang diperolehi melalui eksperimen ke atas kaedah pertama, eksperimen ke atas gabungan dua fitur telah dicadangkan.

KEPUTUSAN

Pengujian, penilaian dan analisis telah dijalankan berdasarkan hasil pengelasan sebutan 30 huruf hijaiyah. Pengukur prestasi yang sesuai telah dipilih bagi menilai prestasi teknik pengelasan yang telah diuji dalam kajian ini. Ketepatan telah dipilih sebagai pengukur prestasi sejajar dengan metrik yang sering digunakan oleh penyelidik lain dalam kajian bidang ini. Peratus ketepatan diperolehi dengan menggunakan persamaan matematik 3 di bawah. Bagi tujuan perbandingan dan penilaian teknik pengelasan yang telah digunakan dalam kajian ini, hasil ujian yang telah dijalankan oleh beberapa kajian lepas telah dipilih.

$$\text{Ketepatan} = \frac{\text{Jumlah ramalan yang betul}}{\text{Jumlah ramalan keseluruhan}} \times 100\%$$

Seksyen ini menerangkan hasil kajian ini ke atas pengelasan sebutan 30 huruf hijaiyah berdasarkan enam teknik pengelasan yang diuji iaitu KNN, J48, ANN, NB, SVM dan RF. Berdasarkan eksperimen ke atas fitur tunggal, Jadual 2 di bawah menunjukkan nilai purata prestasi ketepatan ke atas keenam-enam teknik pengelasan. Bagi fitur MFCC, purata ketepatan pengelasan bagi teknik pengelasan RF adalah yang paling tinggi iaitu 98.29%, diikuti ANN sebanyak 94.67%, J48 pula 92.69%, KNN 91.94%. Manakala SVM pula sebanyak 86.08% dan NB mencapai prestasi ketepatan sebanyak 80.85%.

Manakala bagi fitur *pitch* pula, teknik pengelasan J48 merupakan teknik pengelasan yang terbaik bagi mengelaskan fitur *pitch* berbanding teknik pengelasan yang baik kerana telah mencapai prestasi ketepatan sebanyak 60.72%. Diikuti dengan KNN dan RF sebanyak 59.46%. Seterusnya, NB 56.11%, ANN 54.51%, diikuti SVM, 53.92%.

Seterusnya, teknik pengelasan J48 merupakan teknik pengelasan yang terbaik bagi mengelaskan sebutan 30 huruf hijaiyah berdasarkan fitur keamatan kerana telah mencapai purata pengelasan yang tertinggi sebanyak 67.58%. Prestasi ini diikuti oleh NB sebanyak 61.85%, KNN 61.84%, RF 61.69% seterusnya ANN dan SVM masing-masing 59.75% dan 58.75%.

Akhir sekali, prestasi ketepatan bagi fitur KPS sahaja, teknik pengelasan KNN menghasilkan ketepatan pengelasan sebanyak 89.21% diikuti RF sebanyak 89.13%, J48 73.14% seterusnya ANN, NB dan SVM masing-masing 62.8%, 61.56% dan 60.26%.

JADUAL 2. Hasil eksperimen 1

Teknik Pengelasan	Fitur Pertuturan			
	MFCC	Pitch	Keamatan	KPS
KNN	91.94	59.46	61.74	89.21
J48	92.69	60.72	67.58	73.14
ANN	94.67	54.51	59.75	62.8
NB	80.85	56.11	61.85	61.56
SVM	86.08	53.92	58.75	60.26
RF	98.29	59.46	61.69	89.13

Jadual 3 di bawah pula menunjukkan prestasi pengelasan bagi gabungan dua fitur. Bagi gabungan fitur MFCC & *pitch*, purata pengelasan bagi RF merupakan yang tertinggi iaitu sebanyak 96.76%, diikuti ANN, J48 dan KNN masing-masing 93.05%, 92.13% dan 87.4%. Manakala SVM 85.14% dan NB 79.96%.

Bagi gabungan fitur MFCC dan intensiti, sekali lagi RF merupakan teknik pengelasan terbaik bagi mengelaskan gabungan fitur ini kerana telah mencapai prestasi ketepatan sebanyak 97.31%, diikuti ANN 93.29%, J48 91.86, KNN 90.27%, SVM dan NB masing-masing 86.08% dan 80.76%.

Selain itu, bagi gabungan fitur MFCC dan KPS, purata peratus ketepatan yang tertinggi merupakan bagi teknik pengelasan RF, iaitu sebanyak 95.98%. Seterusnya diikuti dengan ANN, J48, KNN, SVM dan NB, masing-masing 91.84%, 90.85%, 89.19%, 84.76% dan 79.68%.

Di samping itu, bagi gabungan fitur *pitch* dan intensiti, purata peratus ketepatan yang tertinggi merupakan bagi teknik pengelasan KNN, iaitu sebanyak 80.8%. Diikuti RF sebanyak 79.53%, J48 70.89%, ANN 64.58%, NB 63.2% dan SVM 59.05%.

Tambahan lagi, teknik pengelasan RF merupakan teknik pengelasan terbaik bagi mengelaskan sebutan 30 huruf hijaiyah berdasarkan gabungan fitur *pitch* dan KPS yang telah mencapai prestasi ketepatan sebanyak 85.16%. Seterusnya, KNN 80.46%, J48 77.46%, ANN 66.39, diikuti dengan NB dan SVM masing-masing 63.42% dan 60.86%.

Akhir sekali, bagi gabungan fitur KPS dan intensiti, RF memberikan prestasi ketepatan yang baik iaitu sebanyak 85.12%, diikuti KNN 82.22%, J48 79.85%, ANN 68.68%, NB 66.30% dan SVM 63.51%.

JADUAL 3. Hasil eksperimen 2

Teknik Pengelasan	Fitur Pertuturan					
	MFCC & Pitch	MFCC & Keamatan	MFCC & KPS	Pitch & Keamatan	Pitch & KPS	KPS & Keamatan
KNN	87.4	90.27	89.19	80.8	80.46	82.22
J48	91.23	91.86	90.85	70.89	77.46	79.85
ANN	93.05	93.29	91.84	64.58	66.39	68.68
NB	79.96	80.76	79.68	63.2	63.42	66.30
SVM	85.14	86.08	84.76	59.05	60.86	63.51
RF	96.76	97.31	95.98	79.53	85.16	85.12

Berdasarkan Jadual 2 dan Jadual 3 di atas, jelas bahawa prestasi ketepatan fitur tunggal mengatasi prestasi ketepatan gabungan dua fitur. Jadual 3 menunjukkan MFCC sebagai fitur yang mempunyai prestasi ketepatan yang paling baik berbanding fitur yang lain. Hal ini membuktikan MFCC merupakan fitur yang terbaik dan sesuai bagi mengelaskan sebutan huruf-huruf hijaiyah, lebih-lebih lagi dengan penggunaan teknik pengelasan RF.

Manakala berdasarkan Jadual 3 pula, gabungan fitur MFCC dengan fitur intensiti, KPS dan *pitch* memberikan prestasi ketepatan yang baik juga berbanding dengan eksperimen ke atas ketiga-tiga fitur ini secara berasingan. Walaupun terdapat sedikit kemajuan dari segi prestasi ketepatan dengan penggunaan gabungan dua fitur, ianya masih belum mampu untuk mengatasi hasil eksperimen ke atas fitur tunggal dalam mengelaskan sebutan 30 huruf hijaiyah. Selain daripada itu, fitur *pitch* merupakan fitur yang paling kurang sesuai dalam mengelaskan sebutan 30 huruf hijaiyah berdasarkan prestasi ketepatannya dalam mengelaskan sebutan-sebutan tersebut.

Secara purata, berdasarkan Jadual 2 dan 3 yang ditunjukkan di atas, RF merupakan teknik pengelasan yang paling sesuai bagi mengelaskan sebutan 30 huruf hijaiyah. Walaupun begitu, antara keenam-enam teknik pengelasan yang diuji, SVM merupakan teknik pengelasan yang paling kurang memberangsangkan. Oleh itu, sudah terbukti ianya tidak sesuai bagi mengelaskan sebutan 30 huruf hijaiyah. Akhir kata, dapatan kajian ini sejajar dengan dapatan kajian lepas dalam bidang pengelasan sebutan huruf-huruf hijaiyah. Hal ini disokong oleh kajian (Adiwijaya et al., 2017; Elvira, 2018; Wisnu Adi et al., 2018) yang mengatakan MFCC merupakan fitur yang penting dalam mengelaskan sebutan huruf-huruf hijaiyah.

Jadual 4 di bawah pula menunjukkan perbandingan prestasi ketepatan kajian yang dijalankan ini dengan kajian yang dibuat oleh (Elvira, 2018). Kajian (Elvira, 2018) memberi fokus kepada tiga huruf hijaiyah sahaja iaitu ث, س dan ش kerana ketiga-tiga huruf ini mempunyai bunyi yang lebih kurang sama apabila disebut. Perbandingan setara ini telah dibuat berdasarkan fitur MFCC dan teknik pengelasan ANN. Jumlah lipatan K yang digunakan adalah 10. Kajian ini telah berjaya memberikan prestasi ketepatan yang tinggi berbanding dengan kajian (Elvira, 2018) bagi ketiga-tiga huruf tersebut. Faktor yang menyumbang kepada prestasi kajian ini menjadi lebih tinggi berbanding kajian lepas adalah cara rakaman suara diperolehi selain cara pemprosesan data rakaman suara tersebut. Selain itu, cara fitur diekstrak juga mempengaruhi hasil pengelasan yang dibuat dan akhir sekali tetapan parameter teknik pengelasan yang diuji juga memainkan peranan penting.

JADUAL 4. Perbandingan prestasi ketepatan kajian ini dengan kajian yang dijalankan oleh (Elvira, 2018)

Kajian	(Wahyuni 2018)		Ini	
Fitur Pertuturan	MFCC		MFCC	
Teknik Pengelasan	ANN		ANN	
Jumlah Lipatan K	10		10	
Huruf & Prestasi Ketepatan (%)	tsa	91.63	tsa	94.29
	sin	92.38	sin	95.63
	syin	93.26	syin	97.26

KESIMPULAN

Kajian ini telah mengetengahkan isu pengelasan sebutan 30 huruf hijaiyah. Perbandingan antara enam teknik pengelasan berdasarkan pembelajaran mesin. Kajian ini telah menggunakan fitur yang telah diekstrak bagi mengelaskan sebutan 30 huruf hijaiyah. Setiap fitur yang diekstrak membawa nilai yang berbeza-beza dan setiap fitur ini melambangkan setiap sebutan. Hasil kajian ini membuktikan bahawa dengan penggunaan fitur MFCC dan RF sebagai teknik pengelasan mampu memberikan prestasi ketepatan yang baik.

Selain itu, pemilihan fitur juga membantu dalam mengesahkan fitur yang sesuai bagi mengelaskan sebutan 30 huruf hijaiyah dan telah membuktikan bahawa MFCC merupakan fitur yang paling sesuai bagi tujuan kajian ini berbanding fitur-fitur yang lain yang telah diekstrak sejajar dengan dapatan kajian penyelidikan-penyelidik lepas. Ditegaskan bahawa fitur pertuturan memainkan peranan yang penting dalam mengelaskan sebutan 30 huruf hijaiyah.

Kajian ini telah berjaya menambah baik prestasi pengelasan sebutan huruf-huruf hijaiyah berdasarkan tiga kajian lepas ini (Adiwijaya et al., 2017; Elvira, 2018; Wisnu Adi et al., 2018). Tambahan lagi, kajian ini telah mengenalpasti fitur MFCC dan teknik pengelasan RF mampu memberikan prestasi yang baik sekaligus memberikan prestasi ketepatan sebanyak 98.29% berbanding dapatan kajian penyelidikan lain. Dapatan kajian ini menunjukkan bahawa prestasi ketepatan pengelasan boleh ditambah baik dengan pemprosesan isyarat dan pemilihan fitur serta dengan penggunaan pendekatan pembelajaran mesin yang diselia yang bersesuaian. Tambahan lagi, dengan penggunaan teknik pengelasan yang lebih efektif dan terkini lagi canggih seperti pembelajaran dalam boleh dipertimbangkan sebagai cadangan penambahbaikan bagi kajian pada masa akan datang.

Prestasi ketepatan pengelasan sebutan huruf-huruf hijaiyah berjaya ditambah baik dalam kajian ini dengan menggunakan jumlah huruf yang lebih banyak berbanding kajian lepas iaitu 30 huruf. Melalui kajian ini, MFCC berjaya dibuktikan sebagai fitur yang sesuai bagi pengelasan sebutan huruf hijaiyah dan teknik pengelasan RF sebagai teknik pengelasan terbaik.

PENGHARGAAN

Kajian ini dibiayai oleh projek penyelidikan Universiti Kebangsaan Malaysia (UKM) GGPM-2020-029 yang dijalankan di bawah Pusat Penyelidikan Teknologi Perisian dan Pengurusan (SOFTAM), www.ftsm.ukm.my/softam, Fakulti Sains dan Teknologi Maklumat.

RUJUKAN

- Adiwijaya, Aulia, M. N., Mubarak, M. S., Untari Novia, W., & Nhita, F. (2017). A comparative study of MFCC-KNN and LPC-KNN for hijaiyyah letters pronunciation classification system. *2017 5th International Conference on Information and Communication Technology, ICoICT 2017, c*, 2–6. <https://doi.org/10.1109/ICoICT.2017.8074689>
- Al-Sabri, A., Afzan Adam & Fadhilah Rosdi (2018). Automatic detection of Shadda in modern standard Arabic continuous speech. *International Journal on Advanced Science, Engineering and Information Technology*, 8(4-2), 1810–1819. <https://doi.org/10.18517/ijaseit.8.4-2.6813>
- al-Qattan. (1973). *Studi Ilmu-ilmu Qur'an*.
- Andreas Brinch, N., Lars Kai, H., & U K. (2006). Pitch Based Sound Classification. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 3. <https://doi.org/10.1109/icassp.2006.1660772>
- Anon. (2020). *Feature Extraction*. MathWorks. <https://ww2.mathworks.cn/en/discovery/feature-extraction.html>
- Anusuya, M. A., & Katti, S. K. (2009). Speech Recognition by Machine: A Review. (*IJCSIS*) *International Journal of Computer Science and Information Security*, 6(3), 181–205. <http://sites.google.com/site/ijcsis/>
- Atul, G. (2020). *Feature Selection Techniques in Machine Learning*. <https://www.analyticsvidhya.com/blog/2020/10/feature-selection-techniques-in-machine-learning/>
- Azhar Mohd Khairy, Afzan Adam & Mohd Ridzwan Yaakub. (2018). Data Analytics in Malaysian Education System: Revealing The Success of Sijil Pelajaran Malaysia From Ujian Aptitud Sekolah Rendah. *Asia-Pacific Journal of Information Technology & Multimedia*, 07(02), 29–45. <https://doi.org/10.17576/apjitm-2018-0702-03>
- Bachu, R. G., Kopparthi, S., Adapa, B., & Barkana, B. D. (2010). Voiced/unvoiced decision for speech signals based on zero-crossing rate and energy. *Advanced Techniques in Computing Sciences and Software Engineering*, 279–282. https://doi.org/10.1007/978-90-481-3660-5_47
- Ben, L. & Karolina, K. (2021). What is Speech Recognition? Dicapai pada 31 Mac 2022 dari <https://www.techtarget.com/searchcustomerexperience/definition/speech-recognition>
- Chavan, K., & Gawande, U. (2015). Speech recognition in noisy environment, issues and challenges: A review. *Proceedings of the IEEE International Conference on Soft-Computing and Network Security, ICSNS 2015, February 2015*, 23–28. <https://doi.org/10.1109/ICSNS.2015.7292420>
- Cornellius Yudha, W. (2020). *5 SMOTE Techniques for Oversampling your Imbalance Data*. Towards Data Science. <https://towardsdatascience.com/5-smote-techniques-for-oversampling-your-imbalance-data-b8155bdbe2b5>
- Doshi, S. (2018). *Music Feature Extraction in Python*. Towards Data Science. <https://towardsdatascience.com/extract-features-of-music-75a3f9bc265d>
- Dimitris, E. (2021). *Feature Selection for Machine Learning: 3 Categories and 12 Methods*. Towards Data Science. <https://towardsdatascience.com/feature-selection-for-machine-learning-3-categories-and-12-methods-6a4403f86543>
- Elvira Sukma, Wahyuni. (2018). Arabic Speech Recognition using MFCC Feature Extraction and

- ANN Classification. *Proceedings - 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2017, 2018-Janua*, 22–25. <https://doi.org/10.1109/ICITISEE.2017.8285499>
- Ezuana Sukawai & Nazlia Omar. (2020). Pembangunan Korpus Bagi Analisis Sentimen Dalam Bahasa Melayu Secara Separa Selia. *Jurnal Teknologi Maklumat Dan Multimedia Asia-Pasifik*. <https://doi.org/10.17576/apjitm-2020-0901-08>
- Fadhilah Rosdi, Mumtaz Begum Mustafa, Siti Salwah Salim & Nor Azan Mat Zin (2019). Automatic speech intelligibility detection for speakers with speech impairments: The identification of significant speech features. *Sains Malaysiana*, 48(12), 2737–2747. <https://doi.org/10.17576/jsm-2019-4812-15>
- Gabriel, A. (2019). *Feature selection techniques for classification and Python tips for their application*. Towards Data Science. <https://towardsdatascience.com/feature-selection-techniques-for-classification-and-python-tips-for-their-application-10c0ddd7918b>
- Gevaert, W., Tsenov, G., Mladenov, V., & Member, S. (2010). Neural Networks used for Speech Recognition. *Journal of Automatic Control, University of Belgrade*, 20, 1–7. <https://doi.org/10.2298/JAC1001001G>
- Graham, W. (2014). *Loudness*. SLTinfo. <https://www.sltinfo.com/loudness/>
- Hiriyannaiah, S., Srinivas, A. M. D., Shetty, G. K., G.M., S., & Srinivasa, K. G. (2020). A computationally intelligent agent for detecting fake news using generative adversarial networks. In *Hybrid Computational Intelligence* (pp. 69–96). Elsevier. <https://doi.org/10.1016/b978-0-12-818699-2.00004-4>
- Jason, B. (2016). *How to Perform Feature Selection With Machine Learning Data in Weka*. <https://machinelearningmastery.com/perform-feature-selection-machine-learning-data-weka/>
- Jason, B. (2019). *A Gentle Introduction to Imbalanced Classification*. <https://machinelearningmastery.com/what-is-imbalanced-classification/>
- Kaseh Abu-bakar & Muhammad Faiz Abdullah. (2018). Tekanan Perkataan Arab Sebagai Bahasa Asing dalam Kalangan Penutur Melayu. *GEMA Online® Journal of Language Studies*, 18(1). <https://doi.org/10.17576/gema-2018-1801-06>
- Satyam, K. (2021). *Stop using SMOTE to Handle All Your Imbalanced Data*. Towards Data Science. <https://towardsdatascience.com/stop-using-smote-to-handle-all-your-imbalanced-data-34403399d3be>
- Lindasalwa Muda, Mumtaj Begam & Elamvazuthi, I. (2010). Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques. *Journal of Computing*, 2(3), 138–143. <http://arxiv.org/abs/1003.4083>
- Mansour, A., Zulfiqar, A., & Ghulam, M. (2011). Gender Classification with Voice Intensity. *Proceedings - UKSim 5th European Modelling Symposium on Computer Modelling and Simulation, EMS 2011*, 205–209. <https://doi.org/10.1109/EMS.2011.37>
- Mary, W. (2019). *Garbage In, Garbage Out: The Importance of Good Data*. Medium. <https://medium.com/@marybrwolff/garbage-in-garbage-out-the-importance-of-good-data-cel1bb775468e>
- Moawad, A. (2012). Speech Recognition System. In *Global Linguistics* (Issue July 2012, pp. 45–94). <https://doi.org/10.1515/9783110214048.45>
- Nate, G. (2012). *Speech processing pre-emphasis: how does it work?* - Mathematics Stack Exchange. <https://math.stackexchange.com/questions/44216/speech-processing-pre-emphasis-how->

does-it-work

- Nurul Wahidah Arshad, S.N. Abdul Aziz, Faradila Naim, Rohana Abdul Karim, Rosyati Hamid & Nor Farizan Zakaria. (2011). Speech processing for makhray recognition: The design of adaptive filter for noise canceller. *2011 7th International Conference on Information Technology in Asia: Emerging Convergences and Singularity of Forms - Proceedings of CITA'11*. <https://doi.org/10.1109/CITA.2011.5999501>
- Nurul Wahidah Arshad, Suriyazalmi Mohd Sukri, Lailatul Niza Muhammad, Hasan Ahmad, Rosyati Hamid, Faradila Naim & Noor Zirwatul Ahlam Naharuddin. (2013). Makhray Recognition for Al-Quran Recitation using MFCC. *International Journal of Intelligent Information Processing*, 4(2), 45–53. <https://doi.org/10.4156/ijiip.vol4.issue2.5>
- Nilsson, N. J. (2013). *The Quest for Artificial Intelligence*. Cambridge University Press.
- Oksana. (2021). *The Differences Between Audio Formats: MP3, FLAC, WAV, AIFF, M4A & OGG*. <https://www.lalal.ai/blog/difference-between-audio-formats-mp3-flac-wav-aiff-m4a-ogg/>
- Putra, B., Atmaja, B., & Prananto, D., (2012). Developing Speech Recognition System for Quranic Verse Recitation Learning Software. *IJID (International Journal on Informatics for Development)*. 10.14421/ijid.2012.01203
- Rahimi, N. M., Baharudin, H., Zamri, &, & Abstrak, M. (2010). Tahap Sebutan Huruf Konsonan Arab dalam Kalangan Murid Prasekolah (The Level of Arabic Consonants Pronunciation Among Preschool Children). *Jurnal Pendidikan Malaysia*, 35(1), 41–46.
- Rajeev, R., & Rajesh Kumar, D. (2016). Isolated Word Recognition using HMM for Maithili dialect. *2016 International Conference on Signal Processing and Communication, ICSC 2016*, 323–327. <https://doi.org/10.1109/ICSPCom.2016.7980600>
- Shivani, G., & Atul, G. (2019). Dealing with Noise Problem in Machine Learning Data-sets: A Systematic Review. *Elsevier*. <https://doi.org/https://doi.org/10.1016/j.procs.2019.11.146>
- Siti Nor Azimah Sabaruddin & Tengku Intan Zarina Tengku Puji. (2013). Kajian Mengenai Kaedah Pengajaran dan Pembelajaran (P&P) Al-Quran Braille : Suatu Analisa. *Universiti Kebangsaan Malaysia*, 53(9), 1689–1699.
- Schuster (2013). *Sound Intensity and Level*. Boundless Physics. <https://courses.lumenlearning.com/boundless-physics/chapter/sound-intensity-and-level/>
- Sonia, S. & David, P.S. (2013). Performance of Different Classifiers in Speech Recognition. *International Journal of Research in Engineering and Technology*, 02(04), 590–597. <https://doi.org/10.15623/ijret.2013.0204032>
- Swastik, S. (2020). *SMOTE | Overcoming Class Imbalance Problem Using SMOTE*. <https://www.analyticsvidhya.com/blog/2020/10/overcoming-class-imbalance-using-smote-techniques/>
- Tableau. (2021). *Data Cleaning: Definition, Benefits, And How-To*. <https://www.tableau.com/learn/articles/what-is-data-cleaning>
- Teena (2019). How Many Letters in Arabic Alphabet? - Learn Arabic with Teena. Dicapai 5 Oktober 2021 dari <https://learnarabicwithteena.com/arabic-alphabet/arabic-alphabet/>
- ThePhysicsClassroom. (2021). *Physics Tutorial: Intensity and the Decibel Scale*. <https://www.physicsclassroom.com/class/sound/Lesson-2/Intensity-and-the-Decibel-Scale>
- Untari Novia, W., Mubarak, M. S., & Adiwijaya, (2017). A classification of marked hijaiyah letters' pronunciation using hidden Markov model. *AIP Conference Proceedings*. <http://dx.doi.org/10.1063/1.4994439>
- Weenink, D. (2018). *Speech Signal Processing with Praat*.

- Wisnu Adi, P., Adiwijaya, & Untari Novia, W. (2018). Implementation of support vector machine for classification of speech marked hijaiyah letters based on Mel frequency cepstrum coefficient feature extraction. *Journal of Physics: Conference Series*, 971(1). <https://doi.org/10.1088/1742-6596/971/1/012050>
- Xuedong, H., Alex, A., & Hsiao-Wuen, H. (2001). *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development*. Prentice Hall.
- Xuedong, H., & Li, D. (2009). An Overview of Modern Speech Recognition. In *Handbook of Natural Language Processing* (pp. 339–367).
- Ye, Wu., & Rick, R. (2017). *7 Techniques to Handle Imbalanced Data*. <https://www.kdnuggets.com/2017/06/7-techniques-handle-imbalanced-data.html>
- Zitouni, I. (2014). *Natural Language Processing of Semitic Languages*. Springer Science & Business.
- Zbynek, T., & Josef, P. (1999). Speech production based on the mel-frequency cepstral coefficients. *EuroSpeech*, 99, 2335–2338.

PENULIS

Nur Afini Natrah Mohd Ashril merupakan Pelajar Sarjana Sains Komputer (Kecerdasan Buatan) Fakulti Teknologi dan Sains Maklumat (FTSM), Universiti Kebangsaan Malaysia (UKM) dan merupakan pelajar seliaan Dr Afzan Adam dan Dr Dahlila Putri Dahnil.

Afzan Adam (Ph.D) merupakan Pensyarah berpengalaman di Fakulti Teknologi dan Sains Maklumat (FTSM), Universiti Kebangsaan Malaysia (UKM) dengan skop kajian meliputi analitik data, patologi digital, pemprosesan imej serta gabungan kecerdasan buatan, pembelajaran mesin, strategi algoritma dan penyelesaian masalah.

Dahlila Putri Dahnil (Ph.D) merupakan Pensyarah berpengalaman di Fakulti Teknologi dan Sains Maklumat (FTSM), Universiti Kebangsaan Malaysia (UKM) dengan skop kajian meliputi rangkaian sensor tanpa wayar, internet pelbagai perkara (IoT), kejuruteraan perisian masa nyata dan integrasi perisian dan perkakasan.