

BAB II. TINJAUAN PUSTAKA

II.1. Sejarah Kopi

Tanaman Kopi berasal dari Abyssinia atau daerah Afrika yang mencakup wilayah negara Etiopia dan Eritrea. Kemudian dibawa oleh pedagang Arab untuk mempopulerkannya sebagai minuman sampai ke daerah Yaman untuk dijadikan komoditas komersial. Pada abad ke-17 orang-orang Eropa mencoba untuk mengembangkan perkebunan kopi tetapi iklim mereka tidak cocok sehingga membudidayakannya di daerah jajahannya. Upaya orang-orang Eropa berhasil dan menggeser dominasi bangsa Arab dalam memproduksi kopi (Meerbeek et al., 2010). Salah satu pusat produksi kopi di dunia ada di Pulau Jawa yang dikembangkan oleh penjajah Belanda. Pada tahun 1696 Belanda membawa kopi dari Malabar, India, dan pulau Jawa. Mereka membudidayakan tanaman kopi tersebut di Kedawung, Namun upaya ini gagal kerena bencana alam. Pada tahun 1699 Belanda mendatangkan stek pohon kopi dari Malabar ke Indonesia. Pada tahun 1706 kopi yang dihasilkan di Indonesia memiliki kualitas yang sangat baik. Dan Belanda pun memperluas area budidaya kopi ke sampai ke seluruh Indonesia. Pada tahun 1878 seluruh perkebunan kopi yang ada di Indonesia terutama di dataran rendah rusak terserang penyakit. Dan menyisakan tanaman kopi yang berjenis arabika (*Coffea arabica*). Pada tahun 1907 Belanda mendatangkan spesies kopi robusta (*Coffea robusta*) dan berhasil dibudidayakan. Pada saat Indonesia merdeka seluruh perkebunan kopi Belanda yang ada di Indonesia di nasionalisasi. Sejak itu Belanda tidak lagi menjadi pemasok kopi dunia (Asiah, 2022).

II.2. Jenis – Jenis Kopi

Meskipun Indonesia telah lama dikenal mempunyai varietas kopi yang beragam, akan tetapi mayoritas masyarakat hanya membudidayakan dua jenis, yaitu: kopi robusta (*Coffea robusta*) dan kopi arabika (*Coffea arabica*).

II.2.1. Kopi Robusta (*Coffea robusta*)

Kopi robusta merupakan turunan dari sejumlah spesies kopi yang berbeda. Faktanya, kualitas buah kopi lebih rendah dari kopi arabika atau liberika. Varietas kopi ini menyumbang sekitar 30% dari pasar global (Sudarto, 2017). Kopi ini mampu tumbuh dengan baik pada tanah mineral dengan ketinggian 300 – 900 mdpl. Keunggulan dari kopi robusta ini yaitu lebih resisten terhadap serangan hama dan penyakit (Hadi et al., 2014). Menurut (Puspitasari, 2020), Ciri-ciri kopi robusta diantaranya:

1. Aroma lebih *earthy* dan *nutty*
2. Rasanya cukup pahit di bandingkan arabika
3. Teksturnya lebih kasar

4. Kadar gula yang rendah dan kafeinnya tinggi
5. Harga kopi robusta lebih murah dibandingkan dengan kopi arabika

Dapat diketahui tanaman kopi robusta juga tahan terhadap berbagai penyakit oleh lingkungan yang berubah sehingga mempercepat perkembangannya. Karena sifatnya yang lebih unggul kopi berjenis robusta ini banyak di budidayakan di Indonesia (Riastuti et al., 2021). Klasifikasi kopi berdasarkan tingkatan taksonomi, sebagai berikut:

Kingdom : *Plantae*
 Sub Kingdom : *Tracheobionta*
 Super Divisi : *Spermatophyta*
 Divisi : *Magnoliophyta*
 Kelas : *Magnoliopsida*
 Sub Kelas : *Asteridae*
 Ordo : *Rubiales*
 Famili : *Rubiaceae*
 Genus : *Coffea L.*
 Spesies : *Coffea canephora* (Lim, 2013).



(a)



(b)

Gambar 2.1. Visualisasi Bagian Tanaman Coffea canephora (a) Tanaman kopi robusta (b) Biji kopi robusta (Lim, 2013).

II.2.2. Kopi Arabika (*Coffea arabica*)

Kopi arabika adalah turunan dari sejumlah spesies kopi yang berbeda. Faktanya, permintaan kopi arabika lebih tinggi dibandingkan dengan kopi robusta. Varietas kopi ini menyumbang sekitar 70% dari pasar global (Sudarto, 2017). Kopi ini mampu tumbuh dengan baik pada tanah mineral dengan ketinggian lebih dari 1000 mdpl. Memang kopi arabika ini terkenal jika

tanamannya tidak mudah untuk tumbuh dan dirawat tidak seperti tanaman kopi robusta (Hadi et al., 2014). Menurut (Puspitasari, 2020), Ciri-ciri kopi arabika diantaranya:

1. Memiliki aroma yang khas.
2. Bau seperti wangi bunga atau buah
3. Tekstur sedikit lembut dibandingkan kopi robusta
4. Biji kopi nya lebih besar
5. Untuk harga kopi arabika ini 3 lipat lebih mahal dibanding kopi robusta

II.3. Kopi Luwak

Menurut buku Edy Panggabean dengan judul “*The Coffee Smart Book*”, Kopi luwak adalah kopi yang terbuat dari biji kopi yang telah dikonsumsi dan melewati saluran pencernaan luwak, atau lebih tepatnya proses fermentasi biji kopi ini terjadi alami di perut luwak kopi.



Gambar 2.2. (a) Hewan luwak pandan; (b) Hasil kotoran luwak bersama biji kopinya (Uribe et al., 2017).

Sejarah kopi luwak tidak dapat dipisahkan dari sejarah budidaya kopi di Indonesia. Pada awal abad ke-18, Belanda mendirikan perkebunan komersial di koloni mereka di Hindia Belanda, khususnya di Pulau Jawa dan Sumatera. Selama era "Tanam Paksa" atau *Cultuurstelsel* (1830 - 1870), Belanda melarang pekerja perkebunan pribumi memetik buah kopi untuk konsumsi pribadi, tetapi penduduk setempat ingin mencicipi minuman kopi yang terkenal itu. Kemudian para pekerja perkebunan menemukan spesies musang yang suka memakan buah kopi, tetapi hanya daging buahnya yang dicerna, meninggalkan kulit ari dan biji kopi utuh dan tidak tercerna. Kopi luwak kemudian dibuat dengan mengumpulkan biji kopi yang ditemukan di kotoran luwak, mencuci, memanggang, menggiling, dan menyeduh dengan air panas (Raveendran, A., dan Murthy, 2021).

II.4. Komposisi Zat Kimia Pada Kopi

Kandungan kimia pada kopi robusta diantaranya karbohidrat, senyawa nitrogen (protein, kafein, asam amino bebas dan trigonelin), lemak (minyak kopi dan diterpen), mineral, asam dan ester (asam klorogenat dan asam kuinat). Senyawa-senyawa tersebut memiliki peran penting untuk menghasilkan aroma dan rasa pada minuman kopi. Untuk kafein memiliki efek menstimulasi sistem saraf pusat sebagai antagonis reseptor adenosine (Farah, 2012).

II.5. Pengolahan Produk Kopi

Tanaman kopi dapat diperbanyak dengan cara vegetatif menggunakan bagian dari tanaman dan dengan cara generatif menggunakan benih atau biji (Rizwan, 2021). Proses pemanenan buah kopi secara umum dilakukan dengan cara memetik buah yang telah masak berusia sekitar 2,5 – 3 tahun. Untuk mendapatkan hasil yang bermutu tinggi buah kopi harus dipetik dengan keadaan masak penuh. Kopi arabika membutuhkan waktu 6 sampai 8 bulan sejak dari kuncup sampai matang dan untuk kopi robusta hanya membutuhkan 8 – 11 bulan (Duniaji, 2017). Kopi jenis robusta yang ditanam di daerah yang kering biasanya menghasilkan buah pada musim tertentu, sehingga pemanenan dilakukan secara musiman (Widia, W dan Duniaji, 2016).

Biji kopi yang sudah siap untuk diperdagangkan adalah biji kopi kering yang sudah terlepas dari daging buah dan kulitnya. Setelah itu, biji kopi tersebut akan mengalami proses *roasting*, penggilingan, pengemasan hingga diperoleh kopi bubuk yang siap untuk diperjualkan. Proses *roasting* merupakan penyangraian biji kopi yang tergantung pada waktu dan suhu, dan ditandai dengan adanya perubahan kimiawi yang signifikan (Mayrowani, 2013). Menurut (Ruwanto et al., 2019), berdasarkan suhu dan lama penyangraian yang digunakan kopi sangrai dibagi atas tiga golongan yaitu:

1. *Light roast* pada suhu 230°C selama 12 menit.
2. *Medium roast* pada suhu 240°C selama 14 menit.
3. *Dark roast* pada suhu 250°C selama 17 menit.

Pada tahap awal dilakukan pembuangan uap air pada suhu penyangraian 100°C dan berlanjut 180°C pada tahap pirolisisnya. Proses *roasting* berlangsung selama 5 sampai 30 menit (Hećimović et al., 2011).

Proses fermentasi merupakan ciri dari metode pengolahan basah. Fermentasi dilakukan untuk mengeluarkan daging buah yang lengket sehingga mudah lepas saat pencucian dan lebih cepat saat pengeringan. Waktu yang diperlukan untuk membersihkan lendir adalah 36 jam. Ketebalan

lapisan mukus dan konsentrasi enzim yang digunakan juga dapat mempengaruhi sekresi mukus selama fermentasi (Haile and Kang, 2019).

II.6. Adulteran

Adulteran adalah makanan palsu yang dengan sengaja dapat mempengaruhi kualitas makanan. Pemalsuan adalah upaya untuk dengan sengaja mengubah tampilan makanan. Ini dapat dicapai dengan menggabungkan atau mengganti zat yang lebih rendah untuk bahan yang berharga, atau dengan menghilangkan sebagian dari bahan yang berharga sama sekali. Selain itu, istilah "pezina" dapat merujuk pada zat apa pun yang tidak dapat digunakan guna mencemari makanan ataupun yang mengandung zat asing. Pemalsuan makanan tidak terbatas pada penambahan bahan yang disengaja yang menurunkan kualitas makanan. Pencemaran yang dihasilkan selama proses penanaman, pemanenan, penyimpanan, pengolahan, transportasi, dan distribusi. Pemalsuan ini dirancang untuk menaikkan berat dan penampilan makanan untuk memaksimalkan profit (Srivastava, 2015).

Harga kopi luwak arabika jauh lebih tinggi dibandingkan dengan robusta. Pemalsuan pada bahan kopi ini dapat merugikan konsumen maupun produsen. Pemalsuan dalam bentuk serbuk tidak mudah untuk diidentifikasi karena tampilan warna kopi yang sama sehingga memerlukan metode-metode khusus. Pada proses pembuatan kopi bubuk dalam pencampuran biji kopi dengan bahan tambahan bertujuan untuk menambah bobot kopi bubuk yang dihasilkan (Dumadi, 2019). Untuk menentukan keaslian makanan serta deteksi pemalsuan makanan ini merupakan hal yang penting untuk perlindungan kesehatan bagi konsumen. Hal ini dapat terjadi karena bisa saja ada konsumen yang alergi terhadap bahan pemalsu yang ditambahkan ke dalam makanan. Pemalsuan makanan biasanya didorong oleh alasan ekonomi dengan tujuan untuk memperoleh keuntungan yang besar dengan cara mencampur bahan berharga tinggi dengan bahan yang bernilai lebih rendah (Jaswir and Guntarti, 2021)

II.7. Ekstraksi

Ekstrak ialah sediaan padat, kental, atau cair yang didapat melalui ekstraksi simplisia nabati atau hewani dalam lingkungan terkontrol yang bebas dari sinar matahari langsung. Ekstraksi ialah proses pemisahan bahan kimia yang larut dalam air dari bahan aktif yang tidak larut seperti serat, karbohidrat, serta protein. Metode ekstraksi yang paling mudah ialah melunakkan bubuk dalam wadah *inert* yang mengandung pelarut yang sesuai (Mukhirani, 2014). Menurut (Septiana, 2012), pelarut untuk ekstraksi diklasifikasikan menurut tingkat polaritasnya, termasuk:

1. Pelarut polar, atau pelarut dengan tingkat polaritas tinggi, sangat cocok untuk mengekstraksi senyawa polar. Ini adalah pelarut universal meskipun sifatnya polar, pelarut ini mampu mengekstraksi senyawa dengan polaritas rendah. Misalnya, air, metanol, etanol, dan asam asetat.
2. Pelarut semipolar, dibandingkan dengan pelarut polar, mempunyai tingkat kepolaran yang lebih rendah. Contohnya termasuk aseton, etil asetat, dan kloroform.
3. Pelarut non polar berguna untuk mengekstraksi senyawa yang tidak larut dalam pelarut polar, seperti bermacam jenis minyak. Misalnya, heksana, eter, dan sebagainya.

II.8. Fourier Transform Infrared Spectroscopy (FTIR)

Teknik spektroskopi inframerah transformasi *Fourier* digunakan untuk mengumpulkan spektrum inframerah. Teknik spektroskopi inframerah transformasi Fourier (FTIR) banyak digunakan untuk mengidentifikasi senyawa alami dan sintetis. Frekuensi tertentu diserap sementara yang lain ditransmisikan ketika sinyal inframerah dilewatkan melalui sampel senyawa organik. Transisi yang diamati dalam penyerapan inframerah terkait dengan perubahan keadaan vibrasi molekul. Spektrum inframerah dihitung sebagai persentase absorbansi. Spektroskopi inframerah melibatkan interaksi berbagai sistem getaran terkait penyerapan dalam molekul untuk menghasilkan puncak yang sangat khas dalam spektrum. Menurut aplikasi dan instrumentasinya, spektroskopi inframerah diklasifikasikan menjadi tiga jenis radiasi: inframerah dekat, inframerah tengah, dan inframerah jauh. Identitas dua senyawa bisa ditentukan dengan membandingkan spektrum inframerahnya. Ikatan tertentu dalam area sidik jari akan menjadi jauh lebih sulit untuk dibedakan (Hayati et al., 2012).

Kelebihan teknik spektroskopi FTIR antara lain:

1. Dapat dilakukan langsung pada sampel tanpa perlu pemisahan terlebih dahulu.
2. Mampu menganalisis banyak komponen dengan cepat, tidak memerlukan preparasi sampel, serta meminimalkan gangguan selama penentuan suatu senyawa.
3. Untuk membedakan sampel pada berbagai tahap (gas, padat atau cair). Salah satu kelemahannya adalah FTIR sulit untuk difiksirkan secara visual dan langsung karena spektrum serapan molekul yang tumpang tindih (Gad et al., 2013).

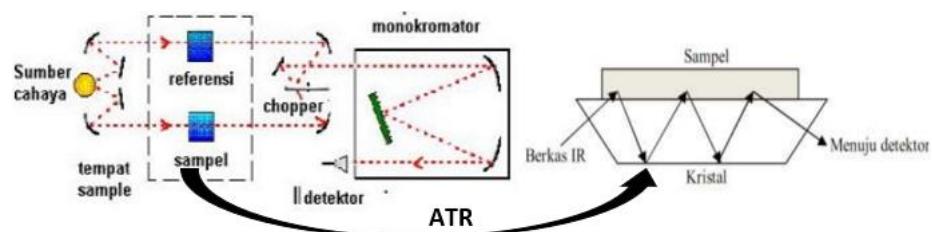
II.8.1. Prinsip Dasar FTIR

Teknik FTIR didasarkan pada prinsip energi inframerah melewati celah ke sampel yang berfungsi mengatur jumlah energi yang dikirim ke sampel. Panjang gelombang inframerah tertentu diserap oleh permukaan sampel, sementara yang lain ditransmisikan melaluinya. Sinar

inframerah melewati detektor yang mentransmisikan sinyal yang diukur ke komputer. Sumber cahaya inframerah ini diatur intensitasnya dengan menggerakkan cermin pantul yang mengarahkan cahaya dari sumber ke sampel. Sinyal detektor berfluktuasi menghasilkan sinyal *interferogram*. Ekspresi yang dihasilkan tidak dapat dipahami dalam bentuk aslinya. Transformasi *Fourier* yang dilakukan secara matematis dengan bantuan komputer, mengubah *interferogram* menjadi spektrum intensitas dan frekuensi (Bunaci et al., 2011). Ketika energi dalam bentuk cahaya inframerah berinteraksi dengan materi dalam bentuk molekul senyawa kompleks membuat molekul tersebut bergetar. Besarnya energi vibrasi setiap komponen molekul bervariasi menurut atom dan kekuatan ikatan yang menghubungkannya menghasilkan rentang frekuensi (Sulistyani, 2018). Berdasarkan tipe pantulan dari sampel dibagi menjadi beberapa tipe, yaitu: *Specular Reflectance* (SR), *Diffused Reflectance* (DR), dan *Attenuated Total Reflectance* (ATR). Teknik ATR merupakan teknik yang paling sering digunakan.

Prinsip ATR yaitu dimana sampel diteteskan pada kristal ZnSe yang terdapat pada spektroskopi FTIR, sinar inframerah akan melewati medium dengan indeks bias yang tinggi ke medium indeks bias rendah yaitu dari kristal ZnSe menuju sampel. Sebagian sinar akan dipantulkan kembali ke sampel pada sudut tertentu sebagian besar gelombang sinar dipantulkan kembali yang disebut dengan *total internal reflection*. Pada kondisi ini sebagian kecil energi akan terlepas dari kristal dan berpindah melewati bagian bawah kristal dalam bentuk gelombang. Dengan adanya pelepasan energi ini akan menyebabkan intensitas sinar yang dipantulkan menjadi bekurang disebut dengan kondisi ATR (Ratnasari, 2016).

II.8.2. Komponen FTIR di lengkapi dengan ATR



Gambar 2.3. Skema alat FTIR yang dilengkapi dengan komponen ATR (Dachriyanus, 2004) dan (Ratnasari, 2016).

1. Sumber cahaya inframerah dari mana sinar berasal. Lampu *tungsten*, *narnst glowers*, dan *glowbars* sering digunakan sebagai sumber cahaya. Monokromator digunakan untuk memilih panjang gelombang dalam spektrofotometer dispersi inframerah.
2. Interferometer, berguna mengatur intensitas sumber cahaya inframerah dengan menggerakkan cermin pemantul yang memantulkan cahaya ke sampel.

3. Sampel, tergantung pada jenis analisisnya, cahaya akan memasuki kompartemen sampel melalui cermin dari permukaan sampel. Tempat sampel ini terdiri dari bahan penyerap inframerah seperti kristal.
4. Detektor bertanggung jawab atas konversi sinyal radiasi inframerah menjadi sinyal listrik. Selain itu, detektor dapat mendeteksi perubahan suhu yang disebabkan oleh pergerakan molekul. Fluktuasi cahaya yang mencapai detektor menghasilkan sinyal interferogram.
5. Komputer, interferogram untuk diubah menjadi spektrum inframerah menggunakan komputer (Bunaci et al., 2011).

II.8.3. Spektrum FTIR

Spektrum FTIR ialah hasil interaksi senyawa kimia dalam matriks sampel yang kompleks. Meskipun komposisi kimia senyawa tersebut belum diketahui secara pasti, spektrum FTIR ini bisa dimanfaatkan untuk membedakan tumbuhan. Spektrum ini dapat digunakan untuk menentukan struktur molekul suatu bahan baku dengan memanfaatkan pita serapan tertentu (Bunaci et al., 2011). Menurut (Rohman, 2014), berikut tabel korelasi yang menyatakan hubungan antara bilangan gelombang dengan gugus gugus fungsi yang bertanggung jawab pada serapan inframerah.

Tabel 2.1. Jenis gugus dan bilangan gelombang (Rohman, 2014).

Gugus	Jenis vibrasi	Bilangan gelombang (cm ⁻¹)	Intensitas
C-H	Alkana (ulur)	3000 – 2850	Kuat
	-CH ₂ - (tekuk)	1450 dan 1375	Medium
	-CH ₂ - (tekuk)	1465	Medium
	Alkena (ulur)	3100 - 3000	Medium
	Alkena (tekuk, keluar bidang)	1000 - 650	Kuat
	Aromatis (ulur)	3150 - 3050	Kuat
	Aromatis (tekuk, keluar bidang)	900 - 690	Kuat
	Alkuna (ulur)	3300	Kuat
	Aldehid	2900 - 2700	Lemah
C-C	Alkana	1200	Sedang
C=C	Alkena	1680 – 1600	Medium-Lemah
	Aromatis	1600 dan 1475	Medium-Lemah
C≡C	Alkuna	2250 - 2100	Medium-Lemah
C=O	Aldehid	1740 – 1720	Kuat
	Keton	1725 – 1705	Kuat
	Asam karboksilat	1725 – 1700	Kuat
	Ester	1750 – 1730	Kuat

	Amida	1680 – 1630	Kuat
	Anhidrida	1810 dan 1760	Kuat
	Asil Klorida	1800	Kuat
C-O	Alkohol, eter, ester, asam karboksilat, anhidrida	1300 - 1000	Kuat
O-H	Fenol Bebas	3650 – 3600	Medium
	Fenol terikat hidrogen	3400 – 3200	Medium
	Asam-asam karboksilat	3400 – 2400	Medium
N-H	Amin primer, amin sekunder, amida (Ulur)	3500 – 3100	Medium
	Amin primer, amin sekunder, amida (Tekuk)	1640 – 1550	Medium-kuat
C-N	Amina	1350 - 1000	Medium-kuat
C=N	Imina dan oksim	1690 – 1640	Medium-kuat
C≡N	Nitril	2260 – 2240	Medium
X=C=Y	Alena, ketena, isosianat, isotiosianat	2270 – 1940	Medium-kuat
N=O	Nitro (R-NO ₂)	1550 dan 1350	Kuat
S-H	Merkaptan	2250	Lemah
	Sulfoksida	1050	Kuat
S=O	Sulfon, sulfonil klorida, sulfat, sulfonamid	1375 - 1300 dan 1350 – 1140	Kuat
	Fluorida	1400 – 1000	Kuat
C-X	Klorida	785 – 540	Kuat
	Bromida, iodide	< 667	Kuat

II.8.4. Analisis Sidik Jari

Analisis sidik jari ialah teknik yang dapat digunakan guna mengevaluasi serta mengontrol kualitas tanaman obat yang mengandung banyak unsur. Analisis ini memberikan informasi mengenai kandungan kimia tanaman obat dalam bentuk spektrogram, kromatogram, dan grafik lain yang diperoleh dengan menggunakan teknik analisis untuk memastikan identitas, kualitas, dan keaslian tanaman obat. Sebab tingginya tingkat variabilitas komponen kimia tanaman obat karena faktor-faktor seperti asal tanaman, proses pengeringan, dan variabel lainnya, maka perlu untuk mengidentifikasi mereka untuk menjaga kepercayaan dalam penelitian klinis dan farmakologis, untuk memahami bioaktivitas, dan potensi efek samping komponen aktif tanaman obat, serta menaikkan pengendalian mutu produk (Borges et al., 2007).

II.9. Analisis Data Kemometrik

Kemometrik ialah subbidang kimia yang banyak menggunakan teknik matematika serta statistik. Teknik ini digunakan untuk memproses, mengevaluasi, dan menginterpretasikan sejumlah besar data dan untuk memilih desain analitik yang menghasilkan hasil eksperimen yang dapat diterima sambil juga memberikan informasi yang relevan. Pola kompleks spektrum inframerah dapat membuat interpretasi langsung menjadi sulit. Teknik kemometrik, seperti analisis multivariat, diperlukan untuk memfasilitasi proses ini. Kemometrik dapat membantu dalam interpretasi data melalui perangkat lunak komputer, menghasilkan hasil yang akurat, sederhana, serta cepat (Gad et al., 2013).

Korelasi antara statistik sampel yang diketahui ditemukan menggunakan metode kemometrik. Teknik ini berpotensi menaikkan kegunaan spektroskopi FTIR secara signifikan untuk analisis tanaman. Dalam rentang tertentu, penggunaan data spektrum ini dapat membantu dalam peningkatan hasil analisis kemometrik. Teknik analisis ini didasarkan pada informasi unik yang terkandung dalam sidik jari. Sebuah properti variabel, seperti aktivitas biologis atau konsentrasi, yang dapat mempengaruhi penampilan kimia sampel (Gad et al., 2013).

II.9.1. Reduksi data PCA (Principal Component Analysis)

Principal Component Analysis (PCA) adalah sebuah teknik untuk membangun variabel-variabel baru yang merupakan kombinasi linear dari variabel-variabel asli (Fadila, 2020). Menurut (Tharwat, 2017), Kelebihan dari PCA adalah menemukan pola dalam data, data dapat dikompresi dengan mengurangi jumlah dimensi tanpa kehilangan banyak informasi. Tujuan utama dari *Principal Component Analysis* (PCA) untuk mengurangi dimensi dataset yang terdiri dari sejumlah besar atribut yang mungkin saling berkorelasi, dengan mempertahankan sebanyak mungkin variansi dataset. Hal ini dicapai dengan mentransformasikan variabel asli yang mungkin berkorelasi menjadi variabel baru yang tidak berkorelasi. Variabel baru ini disebut sebagai *principle component* (PC). Jumlah maksimum *principle component* (PC) adalah sama dengan jumlah variabel aslinya (Anwar, 2018). PC tersebut kemudian diurutkan sehingga variabel-variabel pertama merupakan variabel yang mempertahankan sebagian besar variansi dari keseluruhan variabel aslinya (Tharwat, 2017).

Menurut (Jolliffe, 2002), tahapan untuk melakukan PCA adalah sebagai berikut:

1. Menghitung matriks kovarians C

$$C = \frac{1}{m-1} \sum_{i=1}^m (x_i - \mu)(x_i - \mu)^T \quad (2.1)$$

2. Menghitung nilai eigen (λ) dan vector eigen (v) dari matriks kovarians C

$$Cv = \lambda v \quad (2.2)$$

3. Mengurutkan vector eigen yang berkorespondensi dengan nilai eigen dari yang terbesar hingga yang terkecil. Vector eigen digunakan untuk mentransformasi data menjadi data baru dengan dimensi data yang sama atau lebih kecil
4. Mentransformasi data dengan mengalikan data lama dengan mengalikan matriks eigen yang telah diurutkan.

$$\hat{X} = (v)^T X \quad (2.3)$$

II.9.2. Klasifikasi SVM (Support Vector Machine)

Support vector machines (SVM) pertama kali diusulkan oleh Vladimir Vapnik dan dipresentasikan pertama kali pada tahun 1992 di *Annual Workshop on Computational Learning Theory*. SVM adalah algoritma untuk klasifikasi data linear dan nonlinear (Han, J., Kamber, M. & Pei, 2012). SVM merupakan salah satu algoritma pembelajaran yang dapat menganalisis data dan mengenali pola, dan digunakan untuk melakukan klasifikasi maupun regresi (Kesavaraj and Sukumaran, 2013). *Support vector machine* adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (feature space) berdimensi tinggi.

Konsep dasar dari klasifikasi SVM adalah mencari hyperplane (garis batas) terbaik yang berfungsi sebagai pemisah dua buah kelas data. Pencarian bidang pemisah terbaik inilah yang menjadi inti dari *Support Vector Machine* (Munawarah et al., 2016). *Hyperplane* terbaik dapat ditemukan dengan mengukur margin *hyperplane* tersebut dan mencari titik maksimalnya. Margin adalah jarak antara *hyperplane* dengan data terdekat dari masing-masing kelas. Data yang paling dekat dengan *hyperplane* pada masing-masing kelas tersebut disebut sebagai *support vector* (Yusuf and Priambadha, 2013).

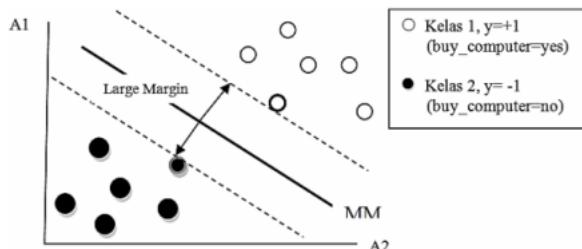
Menurut (Han, J., Kamber, M. & Pei, 2012), langkah awal suatu algoritma SVM adalah pendefenisian persamaan suatu hyperplane pemisah yang dituliskan dengan Persamaan 2.4.

$$W \cdot X + b = 0 \quad (2.4)$$

Di mana w merupakan suatu bobot vektor, yaitu $W = \{w_1, w_2, w_3, \dots, w_n\}$, n adalah jumlah atribut dan b merupakan suatu skalar yang disebut dengan bias. Jika berdasarkan pada atribut A1, A2 dengan permisalan tupel pelatihan $X = (x_1, x_2)$, dimana x_1 dan x_2 merupakan nilai dari atribut A1 dan A2, dan jika b dianggap sebagai suatu bobot tambahan w_0 , maka persamaan suatu *hyperplane* pemisah dapat ditulis seperti pada Persamaan 2.5.

$$w_0 + w_1 x_1 + w_2 x_2 < 0 \quad (2.5)$$

Setelah persamaan dapat didefinisikan, nilai x_1 dan x_2 dapat dimasukkan ke dalam persamaan untuk mencari bobot w_1 , w_2 , dan w_0 atau b. Grafik pemisahan dua kelas data dengan margin maksimum dapat dilihat pada Gambar 2.2.



Gambar 2.4. Pemisahan Dua Kelas Data dengan Margin Maksimum

SVM menemukan *hyperlane* pemisah maksimum, yaitu *hyperlane* yang mempunyai jarak maksimum antara tupel pelatihan terdekat. *Support vector* ditunjukkan dengan batasan tebal pada titik tupel. Dengan demikian, setiap titik yang terletak di atas hyperplane pemisah memenuhi Persamaan 2.6.

$$w_0 + w_1 x_1 + w_2 x_2 > 0 \quad (2.6)$$

Sedangkan, titik yang terletak di bawah hyperlane pemisah memenuhi rumus seperti pada Persamaan 2.7.

$$w_0 + w_1 x_1 + w_2 x_2 < 0 \quad (2.7)$$

Berdasarkan persamaan di atas, bobot dapat disesuaikan sehingga *hyperplanes* yang mendefinisikan sisi dari margin dapat ditulis dalam Persamaan 2.8. dan 2.9.

$$H_1 = w_0 + w_1 x_1 + w_2 x_2 \geq 0, \text{ untuk } y_i = +1 \quad (2.8)$$

$$H_2 = w_0 + w_1 x_1 + w_2 x_2 \leq 0, \text{ untuk } y_i = -1 \quad (2.9)$$

Artinya, setiap tupel yang berada di atas H_1 memiliki kelas +1, dan setiap tupel yang berada di bawah H_2 memiliki kelas -1. Untuk menemukan *Maksimum Margin Hyperplane* (MMH) pada SVM dapat dilakukan dengan menggunakan *Lagrangian formulation* dan kemudian solusi dapat dipecahkan dengan kondisi *Karush - Kuhn - Tucker (KKT)*. Berdasarkan *Lagrangian formulation*, *Maksimum Margin Hyperplane* (MMH) dapat ditulis ulang sebagai suatu batas keputusan (*decision boundary*) seperti pada Persamaan VI.10.

$$d(X^T) = \sum_{i=1}^l y_i a_i X_i X^T + b_0 \quad (2.10)$$

Dimana y_i adalah label kelas *support vector* X_i dan X^T merupakan tupel tes. a_i dan b_0 merupakan parameter numerik yang ditentukan otomatis oleh optimalisasi algoritma SVM dan l merupakan jumlah dari *support vector*.