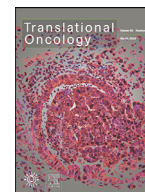




Daftar isi tersedia di [Sains Langsung](#)

Onkologi Translasi

beranda jurnal: www.elsevier.com/locate/tranon



Penemuan biomarker diagnostik kanker paru-paru dini dengan metode pembelajaran mesin

Ying Xie ^{A,1}, Wei- Yu Meng ^{A,1}, Jalankan-Ze Li ^{A,1}, Yu-Wei Wang ^A, Xin Qian ^B, Chang Chan ^B, Zhi-Fang Yu ^B, Xing-Xing Fan ^A, Hu-Dan Pan ^A, Chun Xie ^A, Qi-Biao Wu ^A, Pei-Yu Yan ^A, Liangliu, ^A Yi-Jun Tang ^B, Xiao-Jun Yao ^{A,*}, Mei-Fang Wang ^{B,*}, Elaine Lai-Han Leung ^{A,B,**}

^ALaboratorium Kunci Negara untuk Penelitian Kualitas dalam Pengobatan Tiongkok/Institut Makau untuk Penelitian Terapan dalam Kedokteran dan Kesehatan, Universitas Sains dan Teknologi Makau, Makau (SAR), Tiongkok

^BDepartemen Kedokteran Pernafasan Rumah Sakit Taihe, Universitas Kedokteran Hubei, Hubei, Cina

info artikel

Kata kunci:

Kanker paru-paru
Metabolisme
penanda biologis
Diagnosis dini
Pembelajaran mesin

abstrak

Diagnosis dini terbukti meningkatkan tingkat kelangsungan hidup pasien kanker paru. Ketersediaan skrining berbasis darah dapat meningkatkan penerimaan pasien kanker paru-paru dini. Penelitian kami saat ini berusaha menemukan metabolit plasma pasien Tiongkok sebagai biomarker diagnostik untuk kanker paru-paru. Dalam penelitian ini, kami menggunakan mekanisme interdisipliner perintis, yang pertama kali diterapkan pada kanker paru-paru, untuk mendeteksi biomarker diagnostik kanker paru-paru secara dini dengan menggabungkan metode metabolomik dan pembelajaran mesin. Kami mengumpulkan total 110 pasien kanker paru-paru dan 43 orang sehat dalam penelitian kami. Kadar 61 metabolit plasma berasal dari studi metabolomik yang ditargetkan menggunakan LC-MS/MS. Kombinasi spesifik dari enam biomarker metabolik memungkinkan diskriminasi antara pasien kanker paru stadium I dan individu sehat (AUC = 0,989, Sensitivitas = 98,1%, Spesifisitas = 100,0%). Dan 5 biomarker metabolik yang relatif penting yang dikembangkan oleh algoritma FCBF juga dapat menjadi biomarker skrining yang potensial untuk deteksi dini kanker paru-paru. Naive Bayes direkomendasikan sebagai alat yang dapat dieksploitasi untuk prediksi awal tumor paru-paru. Penelitian ini akan memberikan dukungan kuat terhadap kelayakan skrining berbasis darah, dan menghadirkan alat aplikasi yang lebih akurat, cepat, dan terintegrasi untuk diagnosis dini kanker paru-paru. Metode interdisipliner yang diusulkan dapat disesuaikan dengan kanker lain selain kanker paru-paru.

Perkenalan

Di seluruh dunia, karsinoma paru merupakan penyebab utama kematian akibat kanker dalam beberapa dekade terakhir. Pada bulan Januari 2019, National Central Cancer Registry of China (NCCRC) merilis statistik tumor nasional terbaru dari data registrasi tumor berbasis populasi yang dikumpulkan dari 368 registrasi tumor pada tahun 2015. Menurut laporan ini, kanker paru-paru menempati peringkat teratas dalam hal kejadiannya di antara tumor ganas di dunia. Cina. Terlebih lagi, tingkat kelangsungan hidup 5 tahun pasien tumor paru-paru pun rendah, yaitu 18%. Namun, tingkat kelangsungan hidup dapat meningkat hingga sekitar 55% jika diagnosis dini kanker paru-paru dapat dicapai. Selain itu, dilaporkan bahwa pasien tahap awal yang menerima pengobatan yang tepat dapat memiliki tingkat kelangsungan hidup 5 tahun sekitar 40% [1]. Sayangnya, lebih dari 70% pasien terdiagnosis ketika tumor mereka sudah berkembang ke stadium lanjut, dan sebagian besar dari mereka tidak layak untuk menjalani operasi [2]. Ini sebagian

terkait dengan metode diagnostik tahap awal yang masih belum cukup sensitif dan spesifik. Oleh karena itu, tampaknya merupakan langkah penting untuk mengetahui biomarker diagnostik kanker paru-paru yang meyakinkan dan kuat, khususnya untuk diagnosis perkembangan tumor paru-paru secara dini.

Studi metabolomik telah digunakan untuk mengenali jalur metabolisme dan metabolit yang mengatur perkembangan tumor dan fungsi fisiologis [3-5]. Metabolomik dapat memberikan informasi tentang proses metabolisme seluler yang mendorong tumorigenesis dan perkembangan tumor. Metabolit ini juga dapat membantu untuk membedakan stadium tumor, tipe histologis, dan bahkan respons terhadap pengobatan [6]. Perubahan pola metabolit ini digunakan untuk mengevaluasi karakteristik klinis tumor kolorektal [7], tumor ovarium [8], tumor ginjal [9], tumor mulut [10], dan tumor pankreas [11]. Meski begitu, untuk kanker paru-paru, diperlukan biomarker yang lebih spesifik dan sensitif untuk diungkap melalui analisis metabolik.

*Penulis yang sesuai.

**Penulis koresponden di: State Key Laboratory of Quality Research in Chinese Medicine/Macau Institute for Applied Research in Medicine and Health, Macau University of Science and Technology, Macau (SAR), Tiongkok.

Alamat email: xjyao@must.edu.mo (X.-J. Yao), wmfpps02@hotmail.com (M.-F. Wang), lleung@must.edu.mo (EL-H. Leung).

[†]Para penulis ini memberikan kontribusi yang sama.

<https://doi.org/10.1016/j.tranon.2020.100907>

Diterima pada 4 Juni 2020; Diterima dalam bentuk revisi 21 Agustus 2020; Diterima 25 September 2020

1936-5233/© 2020 Penulis. Diterbitkan oleh Elsevier Inc. Ini adalah artikel akses terbuka di bawah lisensi CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Kecerdasan buatan (AI) adalah kompetensi mesin untuk meniru perilaku manusia, yang sangat mahir dalam menangani data dalam jumlah besar. Pembelajaran mesin adalah penerapan AI, yang memungkinkan sistem komputer dapat dilatih secara otomatis dari pengalaman tanpa diprogram secara eksplisit [12]. Pada dasarnya, pembelajaran mesin berarti belajar dari praktik penggunaan algoritme untuk mengurai data, dan kemudian membuat prediksi atau keputusan tentang situasi masa depan dari kumpulan data baru. [13]. Di bidang kanker, pembelajaran mesin telah digunakan untuk mengeksplorasi model prediksi kelangsungan hidup dan prognosis pada kanker pankreas, kanker kandung kemih, karsinoma nasofaring stadium lanjut, dan kanker payudara. [14-17]. Dalam beberapa kasus, kinerja mereka sebanding dengan kinerja para ahli manusia [18]. Model pembelajaran mesin dapat dipandang sebagai pendekatan merancang model dengan belajar dari pengalaman dan meningkatkan kinerjanya [19]. Model ini bertujuan untuk mengetahui variabel efektif dan hubungan antar variabel tersebut. Selama beberapa tahun terakhir, bidang AI telah beralih dari studi teoretis ke penerapan di dunia nyata [20,21]. Penerapan AI di beberapa domain kini dikaitkan dengan harapan yang besar dan pada saat yang sama terdapat lowongan besar dalam penelitian kanker khususnya kanker paru-paru.

Dalam studi ini, tujuan utama penelitian metabolomik pada kanker paru-paru adalah untuk menemukan biomarker metabolik klinis yang memiliki perubahan representatif antara pasien tumor paru-paru dan individu sehat. Selain itu, kami juga fokus pada biomarker untuk membedakan setiap subtype histologis dan stadium penyakit, terutama untuk stadium awal. Untuk pertama kalinya, berdasarkan fitur metabolit plasma, kami menerapkan pembelajaran mesin untuk mengembangkan model diagnostik kanker paru-paru stadium awal.

Bahan dan metode

Pasien dan kelompok

Sebanyak 110 pasien dan 43 orang sehat di Rumah Sakit Hubei Taihe dilibatkan dalam penelitian ini. Dewan Peninjau Institusi Rumah Sakit Taihe, Universitas Kedokteran Hubei menyetujui penelitian yang melibatkan pasien dan individu sehat. Semua individu telah memberikan persetujuan tertulis sebelum berpartisipasi dalam penyelidikan, dengan izin pengumpulan sampel, penggunaan dan analisis data. Diagnosis akhir dipastikan berdasarkan gejala klinis dan pemeriksaan histopatologi spesimen operasi. Menurut sistem stadium TNM, pasien diklasifikasikan menjadi stadium I ($N=54$) tahap II ($N=31$), tahap III ($N=25$). Berdasarkan klasifikasi tumor WHO [22], tumor telah diklasifikasikan sebagai adenokarsinoma ($N=63$), karsinoma skuamosa ($N=41$) dan tipe histologis lainnya ($N=6$).

Studi metabolomik yang ditargetkan menggunakan LC-MS/MS

Studi metabolomik yang ditargetkan dilakukan dengan metode LC-MS yang dilaporkan sebelumnya [23,24]. Singkatnya, sampel plasma (200 L) dicairkan di atas es, dicampur dengan buffer N-ethylmaleimide PBS (10 mM, 200 L) dan 1000 L metanol yang mengandung 10 ng/ml standar internal (IS) Phed5. Larutan campuran kemudian diinkubasi 20 menit pada suhu -20°C dan disentrifugasi pada 13.000 rpm selama 10 menit pada suhu 4°C . Supernatan dikeringkan dalam aliran nitrogen pada suhu 4°C dan dilarutkan dengan metanol 30% untuk analisis LC-MS.

Pemisahan kromatografi dilakukan dengan Jembatan Waters XTMM kolom analitik BEH C18 (2.5 m, $3.0\times 100\text{mm}$; Waters, Torrance, CA) menggunakan Waters ACQUITY UPLC yang digabungkan dengan spektrometer massa 4000 Q-TRAP. Fase gerak terdiri dari 0,1% air asam format (pelarut A) dan metanol (pelarut B) yang dijalankan dalam program gradien: 0–3,0 menit (0%–1% B); 3,0–10,0 menit (1–3% B); 10,0–14,0 mnt (3–50% B); 14,0–18,0 mnt (50–95% B); 18,0–22,0 menit (95–0% B); diikuti dengan langkah penyeimbangan ulang selama 3 menit. Laju aliran adalah 0,6 ml/menit, dan 10 μL disuntik di LC-MS. Spektrum massa diperoleh dalam mode tegangan ion negatif dan positif untuk electrospray (ESI) dengan parameter berikut: suhu gas, 450°C ; tegangan semprotan ion, $\pm 4500\text{V}$; gas sumber ion 1 (gas nebulizer)

40 psi (N2); gas sumber ion 2 (gas pembantu), 40 psi (N2); gas tirai: 20 psi. Mode Targeted MS/MS (MRM) digunakan dengan energi tumbukan berkisar antara 10 V hingga 40 V. Semua data LC-MS diperoleh oleh Perangkat Lunak Analisis AB (Versi 1.6.2). Intensitas setiap ion dinormalisasi ke area puncak IS sebelum analisis statistik multivariat. Untuk uji metabolomik, analisis komponen utama (PCA) dan proyeksi ortogonal terhadap analisis diskriminasi struktur laten (OPLS-DA) dianalisis dengan perangkat lunak SIMCA-p14 (Umetrics AB, UMEÅ, Swedia) untuk kelompok yang diuji. Variabel disaring berdasarkan nilai VIP terlebih dahulu, dan nilai yang melebihi 1 dianggap memenuhi syarat untuk diskriminasi kelompok. Metabolit yang dipilih selanjutnya dikonfirmasi oleh perbedaan antara kontrol normal dan pengendalian penyakit dengan nilai P kurang dari 0,05. MetaboAnalyst 3.0 digunakan untuk analisis terintegrasi, dampak jalur, dan pengayaan.

Analisis statistik

Semua data statistik dianalisis dengan menggunakan SPSS Statistics 22.0 (SPSS, Chicago, IL, Amerika Serikat), GraphPad Prism 8.0 (GraphPad Software, La Jolla, CA, Amerika Serikat), TBtools v0.6735 [25] dan perangkat lunak Oranye 3.23 [26]. Analisis kurva karakteristik operasi penerima (ROC) dibuat untuk mengevaluasi kinerja diagnostik metabolit. Nilai ϕ -nilai $<0,05$ dianggap signifikan secara statistik.

Metode pembelajaran mesin

Enam teknik pembelajaran mesin K-nearest neighbour (KNN), Naïve Bayes, AdaBoost, Support Vector Machine (SVM), Random Forest, dan Neural Network dengan teknik 10-crossfold digunakan untuk prediksi awal tumor paru berdasarkan metabolisme. fitur biomarker. SVM, yang merupakan algoritma klasifikasi, bermaksud untuk menciptakan batas keputusan antara dua kategori yang memungkinkan prediksi label dari vektor fitur. [27]. K-nearest-neighbor (KNN) adalah pilihan yang disukai ketika hanya ada sedikit pengetahuan sebelumnya tentang data, yang merupakan metode nonparametrik dasar dan sederhana untuk klasifikasi [28]. Random Forest adalah metode pohon ensemble di mana pohon ditanam dengan pemisahan biner secara rekursif dari data yang disensor kanan [29]. Naïve Bayes adalah pengklasifikasi statistik yang digunakan untuk memprediksi probabilitas keanggotaan kelas [30]. Ini menghipotesiskan semua variabel berpartisipasi dalam klasifikasi secara independen dan memberikan hasil prediksi [31]. Neural Network bertujuan untuk mensimulasikan neuron dan otak manusia. Neuron buatan Jaringan Syaraf Tiruan menggunakan fitur masukan tertentu untuk menetapkan bobot matematika yang sesuai yang pada akhirnya mampu memprediksi beberapa objek keluaran [32].

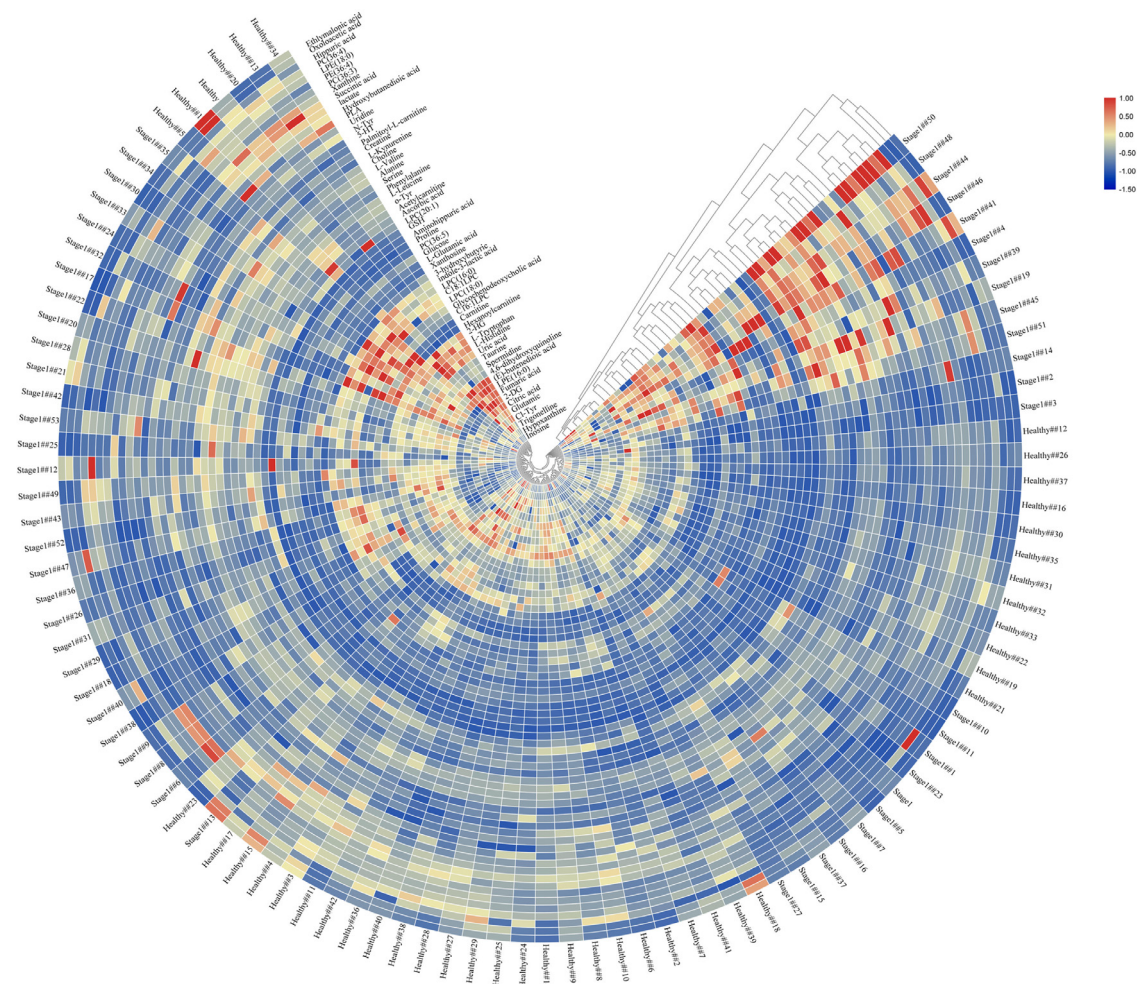
80% sampel termasuk pasien tumor paru stadium I ($N=43$) dan individu sehat ($N=35$) dipilih melalui pengambilan sampel bertingkat dari masing-masing kelompok sebagai set pelatihan untuk melatih model secara seragam. Sisanya 20% sampel termasuk pasien kanker paru stadium I ($N=11$) dan individu sehat ($N=8$) menyusun set tes yang digunakan untuk evaluasi. Set pelatihan digunakan untuk menghasilkan model prediksi yang dapat memprediksi diagnosis set pengujian. Untuk menguji dan membandingkan model, digunakan nilai sensitivitas, spesifisitas, presisi, akurasi klasifikasi, dan AUC (area under the curve) dari masing-masing model untuk mengukur kinerja. Nilai sensitivitas, spesifisitas, dan akurasi klasifikasi diperoleh dari true negative (TN), false negative (FN), true positive (TP), dan false positive (FP). Istilah-istilah berikut adalah informasi penting tentangnya:

TP = Pasien tumor paru yang didiagnosis dengan benar sebagai pasien. FP = Individu sehat salah diidentifikasi sebagai pasien. TN = Individu sehat diidentifikasi dengan benar sebagai sehat. FN = Pasien tumor paru-paru salah diidentifikasi sebagai sehat.

Hasil

Biomarker metabolik untuk mendeteksi tumor paru dini

Penelitian kami bertujuan untuk mengidentifikasi metabolit yang dapat bertindak sebagai biomarker yang menjanjikan untuk membedakan pasien tumor paru-paru dengan individu yang sehat.



Gambar 1. Peta panas yang menggambarkan tingkat biomarker metabolik pasien tumor paru stadium I ($N=54$) dan orang sehat ($N=43$). Pasien tumor paru stadium I dan orang sehat dikelompokkan berdasarkan pengelompokan hierarki tingkat biomarker metabolik.

ual, stadium penyakit serta pola patologi dengan sensitivitas dan spesifisitas tinggi. Pertama, pasien kanker paru stadium I ($N=54$) dan individu sehat ($N=43$) dipisahkan dengan menggunakan pengelompokan hierarki tanpa pengawasan dengan peta panas yang ditunjukkan pada [Gambar 1](#). Melalui uji Mann - Whitney U, 46 biomarker metabolik yang berpengaruh ([Gambar 2A](#) dan Tabel S1) menunjukkan perbedaan yang signifikan secara statistik (*nilai p* < 0,05) di antara 61 metabolit. Peningkatan kadar I-Leucine, I-Valine, serin dan 25 metabolit lainnya diamati pada pasien kanker paru stadium I dibandingkan dengan orang sehat. Selain itu, asam fumarat, asam sitrat, PC (36:4), dan 15 metabolit lainnya mengalami penurunan regulasi pada pasien kanker paru stadium I dibandingkan dengan kontrol yang sehat. 46 metabolit berpengaruh ini merupakan penanda potensial untuk skrining praklinis tumor paru.

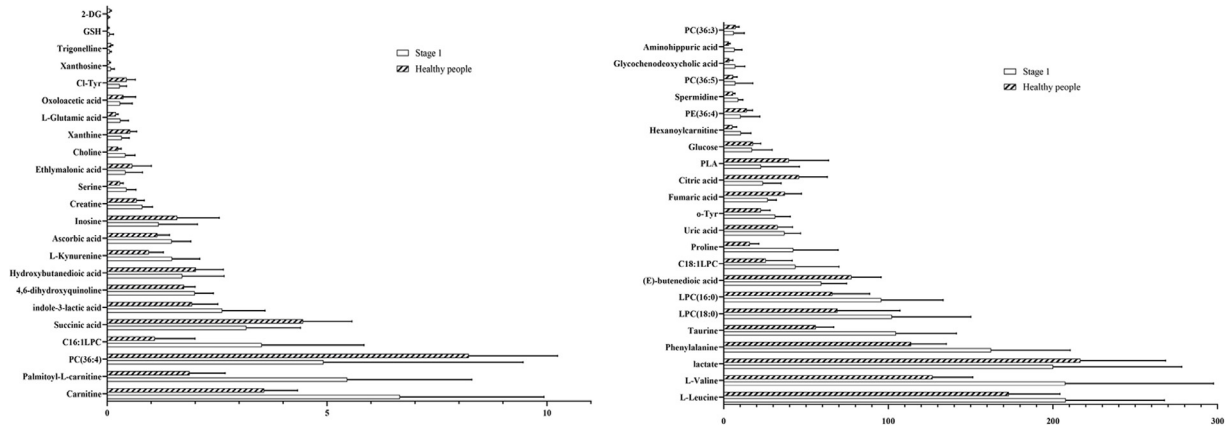
Selanjutnya, 46 biomarker metabolik yang berpengaruh ini diterapkan untuk membuat kurva ROC. Berdasarkan nilai AUC (area di bawah kurva ROC), sensitivitas dan spesifisitas, 10 biomarker metabolik teratas dengan nilai diagnostik lebih tinggi ($AUC > 0,800$) ditunjukkan dalam [Tabel 1](#) dan Gambar. S1. Selain itu, PCA (analisis komponen utama) dengan 10 metabolit ini mengungkapkan pemisahan yang jelas antara pasien tumor paru stadium I dan individu sehat ([Gambar 2B](#) dan [Tabel S2](#)). Pada perbandingan antara pasien tumor paru stadium I dan individu sehat, prolin menunjukkan nilai AUC terbaik yaitu 0,923 (95% CI: 0,871–0,975), dengan sensitivitas 79,6% dan spesifisitas 93,0% pada nilai cut off 24,350. Selain itu, skema kombinasi potensial biomarker metabolik berdasarkan analisis regresi logistik dilakukan untuk meningkatkan sensitivitas dan akurasi diagnostik kanker paru stadium awal. Seperti yang ditunjukkan di [Tabel 1](#), [Tabel S3](#) dan [Gambar 2C](#), kombinasi enam variabel (metabolit) sangat meningkatkan

AUC hingga 0,989 (95% CI: 0,967–1,000, Sensitivitas = 98,1%, Spesifisitas = 100,0%). Metabolit yang digunakan antara lain prolin, l-kynurenine (AUC = 0,825, Sensitivitas = 85,2%, Spesifisitas = 72,1%), spermidine (AUC = 0,890, Sensitivitas = 81,5%, Spesifisitas = 90,7%), asam aminohipurat (AUC = 0,811, Sensitivitas = 68,5%, Spesifisitas = 93,0%), palmitoyl carnitine (AUC = 0,906, Sensitivitas = 74,1%, Spesifisitas = 100,0%) dan taurin (AUC = 0,920, Sensitivitas = 88,9%, Spesifisitas = 95,3%). Hasil ini menunjukkan bahwa 6 biomarker metabolik dapat bertindak sebagai kombinasi yang menjanjikan untuk deteksi dini tumor paru.

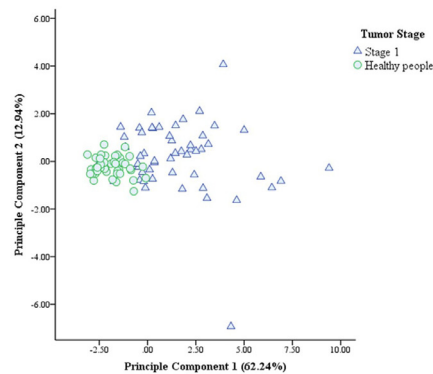
Biomarker metabolik untuk perkembangan penyakit

Selain itu, kami juga tertarik pada perubahan metabolit dalam berbagai tahap. Untuk mengidentifikasi perubahan tingkat metabolit seiring dengan perkembangan stadium tumor, uji Kruskal-Wallis diterapkan. **Gambar 3A** menunjukkan 10 metabolit yang menunjukkan perbedaan signifikan pada tahap I ($N=54$), tahap II ($N=31$), tahap III ($N=25$) pasien tumor paru-paru dan individu sehat ($N=43$). Meskipun terdapat perbedaan yang signifikan secara statistik antara pasien tumor paru-paru dan individu sehat (**Gambar S2 dan Tabel S4**), biomarker metabolik menunjukkan kinerja yang buruk untuk membedakan pasien kanker paru stadium I, II, dan III (**Gambar 3B-D, Tabel S5 dan Tabel S6**). Kurva karakteristik uji non-parametrik dan operasi penerima menunjukkan adanya ekspresi abnormal yang terus menerus pada pasien kanker paru-paru dibandingkan dengan individu sehat.

A



B



Gambar 2. Biomarker metabolik untuk mendeteksi tumor paru dini dan evaluasi tipe histologis yang berbeda. **(A)** 46 biomarker metabolik berpengaruh dengan signifikansi statistik pasien tumor paru stadium I (nilai rata-rata dengan SD). Melalui uji Mann-Whitney U, 46 biomarker metabolik yang berpengaruh menunjukkan perbedaan yang signifikan secara statistik (P -nilai $< 0,05$) di antara 61 metabolit. **(B)** PCA dari 10 biomarker metabolik dalam deteksi dini tumor paru. Ini mengungkapkan pemisahan yang jelas antara pasien tumor paru stadium I dan individu sehat. **(C)** Kurva ROC dari biomarker metabolik dan variasi gabungan dalam deteksi dini tumor paru. Kombinasi enam varian tersebut antara lain prolin, l-kynurenine, spermidine, asam amino-hipurat, palmitoyl-l-carnitine dan taurine. **(D)** Analisis ROC terhadap biomarker metabolik dan variasi gabungan adenokarsinoma ($N=63$) dan karsinoma skuamosa ($N=41$) pasien. Kombinasi empat varian antara lain hipoxantin, l-kynurenine, prolin dan Karnitin. SD, standar deviasi. PCA, analisis komponen utama. ROC, karakteristik pengoperasian penerima.

Biomarker metabolik untuk evaluasi tipe histologis yang berbeda

Untuk prediksi tipe histologis kanker paru-paru, khususnya perbedaan antara karsinoma skuamosa dan adenokarsinoma, ini merupakan persyaratan diagnostik yang signifikan dalam praktik klinis. Beberapa studi klinis telah menunjukkan bahwa jenis histologis tumor berbeda dalam toksisitas dan kemanjuran pengobatan [33]. Identifikasi jenis histologis tumor akan membantu untuk meningkatkan efisiensi pengobatan. Kami menerapkan metabolit dari adenokarsinoma ($N=63$) dan karsinoma skuamosa ($N=41$) pasien membuat kurva ROC dan melakukan uji Mann-Whitney U. Hanya ada 2 metabolit berpengaruh yang diidentifikasi antara adenokarsinoma dan karsinoma skuamosa, termasuk asam hipurat ($nilai\ p=0,029$) dan hipoksantin ($nilai\ p=0,017$). Dalam analisis ROC (Gambar S3), hipoksantin menunjukkan nilai AUC 0,639 (95% CI: 0,531–0,746), dengan sensitivitas 69,8% dan spesifisitas 56,1% pada nilai batas 0,092. Dan asam hipurat menunjukkan nilai AUC 0,628 (95% CI: 0,519–0,737), dengan sensitivitas 49,2% dan spesifisitas 77,5% pada nilai cut off 2,620. Seperti yang ditunjukkan di [Gambar 2D](#) dan [Tabel 1](#), kombinasi empat varian meningkatkan AUC menjadi 0,740 (95% CI: 0,644–0,837,

sensitivitas = 58,7%, spesifisitas = 78,0%), termasuk hipoksantin, l-kynurenine (AUC = 0,423, sensitivitas = 77,8%, spesifisitas = 24,4%), prolin (AUC = 0,580, sensitivitas = 54,0%, spesifisitas = 65,9%) dan Karnitin (AUC = 0,536, sensitivitas = 38,1%, spesifisitas = 75,6%). Hasil ini menunjukkan bahwa biomarker metabolik menunjukkan kinerja yang buruk dalam membedakan tipe histologis tumor paru yang berbeda dalam penelitian kami, karena semua nilai AUC $< 0,800$ dengan sensitivitas dan spesifisitas buruk.

Pemanfaatan metode pembelajaran mesin

Untuk mengembangkan model prediksi tumor paru-paru awal, kami mempertimbangkan enam teknik pembelajaran mesin: K-nearest-neighbor (KNN), Naïve Bayes, AdaBoost, Support Vector Machine (SVM), Random Forest, dan Neural Network ([Gambar 4A](#)). Kumpulan pelatihan pasien tumor paru stadium I ($N=43$) dan individu sehat ($N=35$) digunakan untuk mengembangkan model pembelajaran mesin berdasarkan fitur biomarker metabolik. Dengan menggunakan platform Orange 3.23, terdapat 61 jenis metabolit yang digunakan sebagai fitur untuk mengembangkan model pembelajaran mesin. Untuk menentukan model mana yang memberikan prediksi paling tepat pada data metabolik tumor paru, sensitivitas,

Tabel 1

Analisis ROC dari biomarker metabolik dan variasi gabungan.

	AUC	Std. kesalahan	Interval kepercayaan 95% asimtotik		Pemotongan optimal	Kepekaan	Kekhususan	Indeks Youden
			Batas bawah	Batas atas				
Analisis ROC terhadap biomarker metabolik dan variasi gabungan dalam deteksi dini tumor paru. L-								
Kynurenine	0,825	0,043	0,740	0,909	0,975	85,2%	72,1%	0,573
Prolin	0,923	0,026	0,871	0,975	24.350	79,6%	93,0%	0,727
Spermidin	0,890	0,035	0,821	0,958	7.195	81,5%	90,7%	0,722
Asam amino-hipurat	0,811	0,045	0,722	0,900	4.035	68,5%	93,0%	0,615
Palmitoil-aku-karnitin	0,906	0,032	0,843	0,969	3.655	74,1%	100,0%	0,741
Taurin	0,920	0,032	0,856	0,983	71.300	88,9%	95,3%	0,842
Fenilalanin	0,848	0,038	0,774	0,922	125.500	79,6%	76,7%	0,564
L-Valin	0,876	0,036	0,806	0,946	167.000	68,5%	95,3%	0,639
o-Tyr	0,822	0,043	0,738	0,906	24.650	83,3%	72,1%	0,554
karnitin	0,848	0,040	0,769	0,926	4.680	72,2%	93,0%	0,652
Kombinasi dua	0,933	0,028	0,878	0,978	0,337	85,2%	93,0%	0,782
Kombinasi ketiganya	0,968	0,019	0,931	1.000	- 0,147	94,4%	97,7%	0,921
Kombinasi enam	0,989	0,011	0,967	1.000	- 0,102	98,1%	100,0%	0,981
Analisis ROC terhadap biomarker metabolik dan variasi gabungan pasien adenokarsinoma dan karsinoma skuamosa. L-								
Kynurenine	0,423	0,060	0,306	0,540	1.050	77,8%	24,4%	0,022
Prolin	0,580	0,057	0,469	0,692	35.150	54,0%	65,9%	0,198
karnitin	0,536	0,058	0,422	0,650	6.835	38,1%	75,6%	0,137
Hipoksantin	0,639	0,055	0,531	0,746	0,092	69,8%	56,1%	0,259
Asam hipurat	0,628	0,056	0,519	0,737	2.620	49,2%	77,5%	0,267
Kombinasi empat	0,740	0,049	0,644	0,837	0,556	58,7%	78,0%	0,368

Singkatan: ROC, karakteristik pengoperasian penerima; AUC, area di bawah kurva.

Meja 2

Model pembelajaran mesin digunakan untuk deteksi dini tumor paru berdasarkan fitur biomarker metabolik.

	dll	FP	TN	FN	Akurasi klasifikasi	Kepekaan	Kekhususan	AUC	Presisi
Pelatihan KNN	38	0	35	5	0,936	0,884	1.000	1.000	0,944
	43	2	33	0	0,974	1.000	0,943	1.000	0,975
	Hutan Acak	41	0	35	2	0,974	1.000	1.000	0,976
	Jaringan syaraf	43	0	35	0	1.000	1.000	1.000	1.000
	Bayes Naif	43	0	35	0	1.000	1.000	1.000	1.000
Tes	AdaBoost	38	4	31	5	0,885	0,884	0,885	0,885
	KNN	9	0	8	2	0,895	0,818	1.000	0,916
	SVM	10	0	8	1	0,947	0,909	1.000	0,953
	Hutan Acak	11	2	6	0	0,895	1.000	0,750	0,911
	Jaringan syaraf	10	0	8	1	0,947	0,909	1.000	0,953
mengatur	Bayes Naif	11	0	8	0	1.000	1.000	1.000	1.000
	AdaBoost	4	0	8	7	0,632	0,364	1.000	0,682

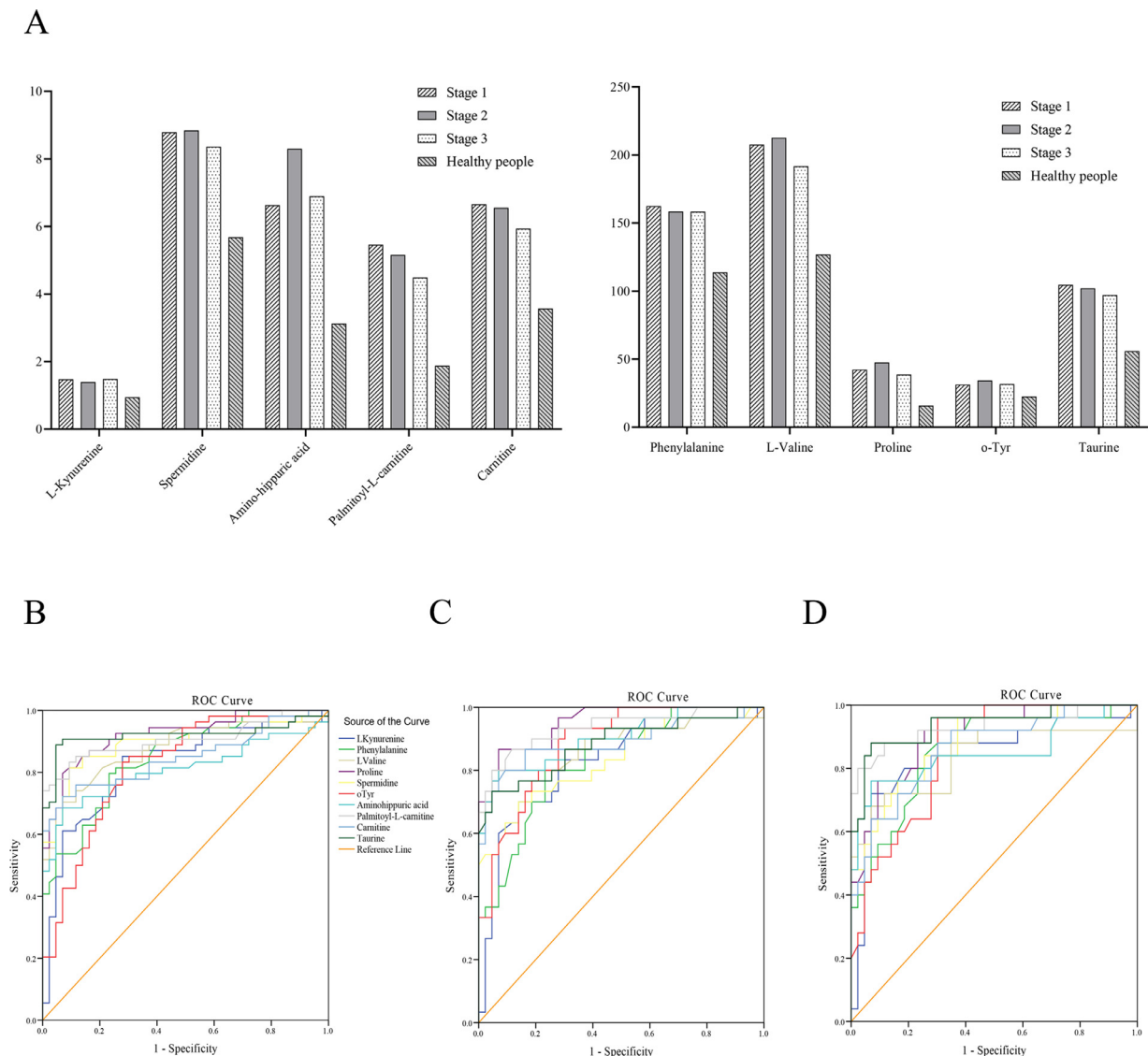
Singkatan: AdaBoost, Adaptive Boosting; SVM, mendukung mesin vektor; KNN, k-tetangga terdekat; TN, benar-benar negatif; FN, negatif palsu; TP, benar-benar positif; FP, positif palsu; AUC, area di bawah kurva.

spesifisitas, presisi, akurasi klasifikasi, dan nilai AUC dari enam model pembelajaran mesin dinilai (Metode tambahan). Dengan nilai terbaik di setiap evaluasi disorot **Meja 2**. Pada set pelatihan, Naïve Bayes dan Neural Network menunjukkan hasil yang lebih baik dibandingkan dengan teknik lainnya (KNN, AdaBoost, SVM, dan Random Forest). Presisi, akurasi klasifikasi, spesifisitas, sensitivitas dan nilai AUC Naïve Bayes dan Neural Network adalah 100,0%. Teknik pembelajaran mesin AdaBoost menunjukkan kinerja yang buruk (presisi =0,885, akurasi klasifikasi=0,885, spesifisitas=0,886, sensitivitas=0,884, dan AUC=0,885).

Selanjutnya, kami menggunakan model pembelajaran mesin diagnostik terlatih untuk mengklasifikasikan set pengujian yang terdiri dari pasien kanker paru-paru stadium I (N=11) dan individu sehat (N=8) untuk mengevaluasi kinerjanya. Seperti yang ditunjukkan di **Meja 2**, model Naïve Bayes menunjukkan kinerja terbaik pada semua parameter evaluasi. Spesifisitas Naïve Bayes, Neural Network, KNN, AdaBoost, dan SVM adalah 1.000 yang berarti daya prediksi yang baik terhadap individu sehat. Untuk kualitas prediksi secara keseluruhan, AUC Naïve Bayes, Neural Network, KNN, Random Forest, dan SVM adalah 1.000. Sensitivitas Naïve Bayes dan Random Forest sebesar 1.000, SVM dan Neural Network sebesar 0.909 yang berarti skala FN (false negative) kecil. Dalam situasi medis, skala FN lebih penting daripada FP (positif palsu) [34]. Akibatnya, keempat model ini (Naïf

Bayes, Random Forest, SVM dan Neural Network) sesuai untuk diagnosis tumor paru dini. Pada akurasi klasifikasi, model Naïve Bayes memiliki tingkat akurasi 1.000 dengan model SVM dan Neural Network sebesar 0.947. Ketiga model tersebut memiliki akurasi yang cukup sehingga dapat digunakan untuk aplikasi medis. Naïve Bayes adalah pengklasifikasi probabilistik sederhana berdasarkan penerapan teorema Bayes dengan asumsi independensi dan normalitas yang kuat antar variabel [35]. Ini adalah salah satu algoritme pembelajaran mesin paling valid dengan independensi yang kuat dan asumsi normalitas antar fitur, yang telah banyak digunakan untuk prediksi [36]. Mengingat hal di atas, penelitian kami membuktikan bahwa Naïve Bayes adalah model terbaik dengan tingkat sensitivitas, spesifisitas, dan akurasi tertinggi. Oleh karena itu, Naïve Bayes direkomendasikan sebagai alat yang dapat dieksploitasi untuk prediksi dini tumor paru-paru.

Algoritma Fast Correlation-Based Filter (FCBF) adalah metode terawasi yang didasarkan pada teori informasi [37]. Dibutuhkan identifikasi fitur-fitur yang berkorelasi untuk klasifikasi dan menghilangkan fitur-fitur yang berlebihan [38]. Berdasarkan ketidakpastian simetris (SU), FCBF mengurutkan fitur-fitur dalam urutan korelasi menurun dan membuang fitur-fitur berlebihan yang kurang berkorelasi [39]. Akibatnya, subset fitur berkorelasi dan non-redundan yang optimal diperoleh. Ini dapat memaksimalkan potensi diagnostik dari informasi yang dapat diekstraksi. Pentingnya relatif dari fitur biomarker metabolik dijelaskan



Gambar 3. Biomarker metabolik berubah seiring kemajuan stadium tumor. (A) Hal ini menunjukkan kadar 10 metabolit yang menunjukkan perbedaan yang signifikan pada tahap I ($N=54$), tahap II ($N=31$), tahap III ($N=25$) pasien tumor paru-paru dan individu sehat ($N=43$). (B) Kurva ROC biomarker metabolik tahap I ($N=54$) pasien tumor paru-paru. (C) Kurva ROC biomarker metabolik tahap II ($N=31$) pasien tumor paru-paru. (D) Kurva ROC biomarker metabolik stadium III ($N=25$) pasien tumor paru-paru. ROC, karakteristik pengoperasian penerima.

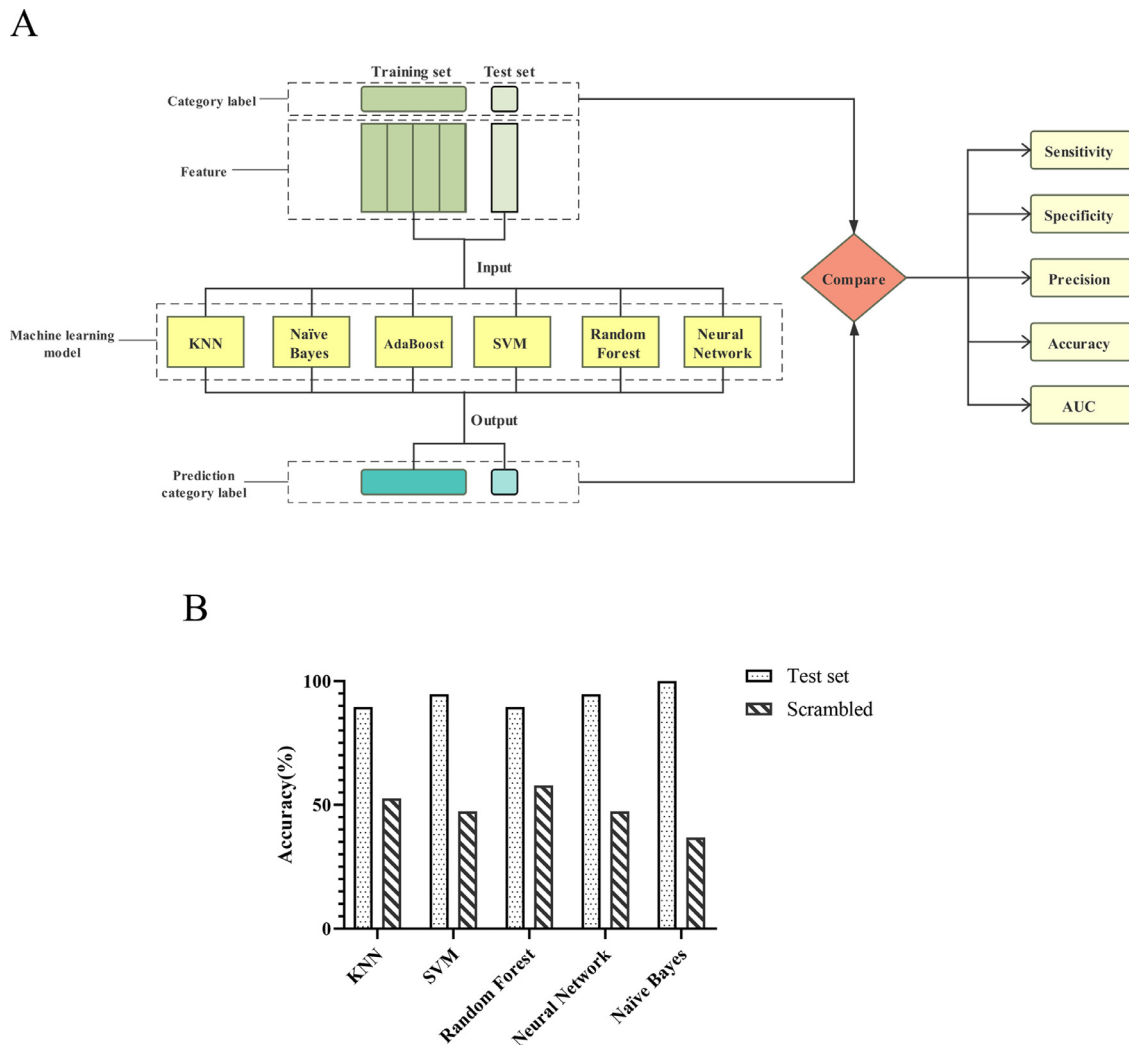
dikembangkan menggunakan algoritma Fast Correlation-Based Filter (FCBF). Semua fitur metabolit diberi peringkat dan diberi skor sesuai dengan kemampuannya dalam membedakan label klasifikasi suatu objek (Tabel S7). Menurut pemeringkatan tersebut, 8 biomarker metabolik teratas masing-masing digunakan untuk mengembangkan model pembelajaran mesin. Tabel S8 dan Gambar S4 menunjukkan bahwa nilai AUC berubah seiring dengan banyaknya variasi. Ketika 5 varian teratas termasuk taurin, Palmitoyl-L-karnitin, prolin, 2-DG, dan PE (36:4) digunakan sebagai fitur metabolit, model prediksi menunjukkan kinerja luar biasa yang serupa dengan model sebelumnya. Oleh karena itu, 5 biomarker metabolik ini dapat menjadi kandidat potensial untuk skrining praklinis kanker paru-paru. Kemudian berdasarkan hasil analisis regresi logistik, kombinasi tiga varian teratas antara lain taurin, Palmitoyl-L-karnitin, dan prolin menunjukkan nilai AUC sebesar 0,968 (95% CI: 0,931–1,000), dengan sensitivitas 94,4% dan spesifisitas 97,7% pada analisis klasik (Tabel 1).

Untuk memvalidasi kinerja diagnosis model pembelajaran mesin dan menunjukkan kekhususan fitur biomarker metabolik pasien kanker paru-paru tahap awal yang ditemukan dalam penelitian kami, kami melakukan eksperimen kontrol [40]. Kami membuat set pelatihan acak yang diberi label pelatihan dengan benar.

set pelatihan diganti dengan set pelatihan yang labelnya ditetapkan secara acak. Seperti yang diharapkan, keakuratan model pembelajaran mesin kami menurun drastis, yang setara dengan memilih secara acak, tidak menunjukkan nilai prediktif, dan memvalidasi tanda prediktif spesifik dari fitur biomarker metabolik (Gambar 4B).

Diskusi

Kanker paru-paru merupakan penyebab utama kematian akibat kanker di seluruh dunia, sehingga diagnosis dini dapat meningkatkan tingkat kelangsungan hidup. Namun, orang yang berisiko tinggi umumnya direkomendasikan untuk melakukan pemeriksaan radiologi tahunan dengan menggunakan tomografi komputer dosis rendah (LDCT). [41], yang juga merupakan satu-satunya cara deteksi kanker paru-paru secara klinis saat ini. Karena biaya yang signifikan dan tingkat penemuan palsu yang tinggi [42], pemenuhan skrining CT kurang memuaskan. Oleh karena itu, ketersediaan skrining berbasis darah dapat meningkatkan penyerapan pasien kanker paru-paru, termasuk deteksi biomarker metabolik plasma. Penelitian kami saat ini berusaha untuk menemukan metabolit plasma pasien Tiongkok sebagai biomarker prediktif untuk diagnosis kanker paru-paru. Dalam pekerjaan ini, kami menggunakan interdisipliner perintis



Gambar 4. Pembelajaran mesin diterapkan untuk mengembangkan model diagnostik kanker paru-paru stadium awal. **(A)** Aplikasi pembelajaran mesin membangun model prediksi tumor paru-paru awal. **(B)** Untuk memvalidasi kinerja diagnosis model pembelajaran mesin dan menunjukkan kekhususan fitur biomarker metabolik pasien kanker paru-paru tahap awal yang ditemukan dalam penelitian kami, kami membuat rangkaian acak yang tidak menunjukkan nilai prediktif. AUC, area di bawah kurva. AdaBoost, Peningkatan Adaptif. SVM, mendukung mesin vektor. KNN, k-tetangga terdekat.

mekanisme, yang pertama kali diterapkan pada kanker paru-paru, untuk mendeteksi biomarker diagnostik kanker paru-paru secara dini dengan menggabungkan metode metabolomik dan pembelajaran mesin. Teknologi metabolomik dan metode pembelajaran mesin yang sangat sensitif dan akurat membawa hal baru pada pekerjaan kami dibandingkan dengan penelitian sebelumnya yang menggunakan miRNA sebagai biomarker untuk diagnosis dini kanker paru-paru.

Penelitian utama sebelumnya mengenai kanker payudara, kanker prostat, dan kanker lainnya telah menyaring miRNA sebagai biomarker untuk membedakan pasien kanker dan subtipe tumor yang berbeda.[43]. Namun demikian, skrining miRNA memiliki keterbatasan, seperti biaya tinggi dan monopoli teknis [44]. Kemudian tes klinis biasa antigen darah CEA, CA125, SCC, dll masih menemui masalah sensitivitas rendah dan akurasi rendah[45]. Pada saat yang sama, spesimen biopsi kecil berarti trauma dan negatif palsu sebagai lokasi bagian yang terbatas[46]. Dalam penelitian kami saat ini, kami fokus pada pencarian biomarker metabolik sebagai biomarker diagnostik kanker paru tahap awal. Di sini, kami mengidentifikasi serangkaian metabolit baru dengan menentukan profil plasmanya pada pasien sebelum diagnosis klinis tumor paru-paru. Dari 61 jenis metabolit, kami menemukan 10 biomarker metabolik yang berpotensi sebagai biomarker yang menjanjikan untuk deteksi dini tumor paru. Khususnya, kombinasi enam varian secara signifikan meningkatkan AUC menjadi 0,989 dengan sensitivitas 98,1% dan spesifisitas 100,0%, yang belum dilaporkan sejauh ini. Berdasarkan hasil ini, kami merekomendasikan kombinasi spesifik dari enam metabolit ini sebagai biomarker skrining

deteksi dini tumor paru. Kami percaya bahwa temuan ini memungkinkan kami mengembangkan penerapan yang spesifik, sensitif, dan invasif minimal untuk pencegahan dan prediksi tumor paru-paru secara dini.

Hingga saat ini, pendekatan pembelajaran mesin, yang memberikan alternatif menjanjikan terhadap metode analisis data klasik, telah digunakan dalam berbagai aplikasi biomedis, termasuk penemuan obat.[47], dan pengembangan biomarker[48]. Daripada jenis analisis data konvensional yang memerlukan kesadaran akan ketergantungan biologis, pembelajaran mesin dapat meningkatkan kemampuan diagnostik melalui data yang berlimpah dan berkualitas tinggi. Sebuah tim mengumpulkan 23 item data demografi dan parameter terkait tumor dari 102 pasien kanker serviks yang telah menjalani histerektomi radikal untuk pengobatan, dan menyelidiki beragam model pembelajaran mesin untuk memprediksi tingkat kelangsungan hidup 5 tahun pada pasien.[49]. Kelompok lain memanfaatkan temuan kolposkopi 2.267 wanita dan biomarker human papillomavirus (HPV) untuk mengembangkan sistem penilaian pendukung keputusan klinis menggunakan jaringan saraf tiruan untuk pasien neoplasia intraepitel serviks, yang menunjukkan prediksi ANN dengan akurasi lebih tinggi dibandingkan sitologi dengan atau tanpa HPV. tes[50]. Selain itu, telah dibuktikan kemungkinan untuk menggali panel mikroRNA serum sebagai biomarker potensial untuk mendeteksi kanker lambung dengan pembelajaran mesin.[51]. Enam jenis teknik pembelajaran mesin digunakan untuk memilih tiga biomarker (miR-21-5p, miR-29c-3p, dan miR-22-3p) dari studi profil miRNA yang dipublikasikan (GSE23739). Di sini, kami

pertama kali menggunakan biomarker metabolik sebagai fitur pembelajaran mesin untuk diagnosis kanker paru-paru, dan hasil yang diperoleh dianalisis dan didiskusikan. 5 biomarker metabolik yang relatif penting (taurin, Palmitoyl-carnitine, proline, PE (36:4) dan 2-DG), yang dikembangkan oleh algoritma FCBF, dapat menjadi kandidat potensial untuk skrining pra-klinis kanker paru-paru. Dalam penelitian ini, beberapa model pembelajaran mesin diterapkan dan dibandingkan, yang dievaluasi melalui set pengujian dan eksperimen kontrol ([Gambar 4](#) [PitaMeja 2](#)). Hal ini menghasilkan nilai penilaian terbesar pada Naïve Bayes, namun juga menghasilkan nilai penilaian yang layak pada Neural Network, dan SVM. Sebagai model skrining mandiri, sensitivitas tinggi diinginkan untuk meminimalkan hasil positif palsu. Seperti yang ditunjukkan di [Meja 2](#), model Naïve Bayes, Random Forest, Neural Network, dan SVM dengan sensitivitas tinggi memiliki potensi yang dapat diandalkan dan stabil untuk prediksi tumor paru dini. Selain itu, ambang batas spesifisitas yang ditetapkan dalam set pelatihan memiliki kinerja serupa ketika diterapkan pada set pengujian, yang menunjukkan bahwa model telah dikalibrasi dengan baik. Ambang batas spesifisitas model Naïve Bayes, Neural Network, dan SVM yang ditentukan dalam set pelatihan mencapai spesifisitas serupa di set pengujian. Mengingat penelitian kami di atas, Naïve Bayes, Neural Network, dan SVM, berdasarkan fitur biomarker metabolik, mungkin kondusif untuk diagnosis tumor paru dini. Hasil ini memberikan dukungan kuat terhadap kelayakan skrining berbasis darah berdasarkan teknologi metabolomik dan pembelajaran mesin untuk diagnosis dini kanker paru-paru.

Perkembangan kanker sangat terkait dengan metabolisme sel, dan metabolisme yang tidak teratur kondusif terhadap perkembangan dan inisiasi tumor [\[52\]](#). Sel kanker mengubah jalur metabolismenya, yang meminta enzim spesifik untuk mengkatalisis reaksi biokimia, untuk memenuhi kebutuhan reproduksi dan pembelahan sel yang cepat. Metabolit, termasuk biomarker metabolik untuk diagnosis kanker paru-paru yang ditemukan dalam penelitian kami, memiliki fungsi yang berkaitan dengan tumorigenesis dan perkembangan tumor. Proline dehydrogenase (PRODH), yang mengkatalisis langkah pertama degradasi prolin, diaktifkan oleh helikase spesifik limfoid (LSH) untuk menurunkan kadar prolin [\[53\]](#). Katabolisme prolin terkait PRODH telah terbukti dapat meningkatkan kelangsungan hidup tumor melalui autophagy yang diinduksi ROS atau produksi ATP, atau sebagai penekan tumor untuk memulai apoptosis yang dimediasi ROS tergantung pada lingkungan mikro tumor [\[54,55\]](#). Sebuah penelitian terbaru dari sebuah tim menemukan bahwa PRODH mendorong tumorigenesis kanker paru-paru dengan memunculkan ekspresi IKK gen inflamasi yang bergantung pada gen dan transisi epitel ke mesenkim (EMT) [\[56\]](#). Tingkat l-kynurenine yang tinggi dapat menyediakan lingkungan mikro untuk pertumbuhan tumor paru-paru melalui inisiasi apoptosis sel T, menghambat proliferasi sel T dan menyebabkan toleransi imun [\[57,58\]](#). Spermine N1-acetyltransferase (SSAT) adalah protein penting yang terlibat dalam homeostasis dan sintesis spermidine [\[59\]](#). Spermidin/SSAT adalah langkah pembatas laju dalam katabolisme poliamina, yang memainkan peran khusus dalam menjaga potensi membran dan mengatur volume sel [\[60,61\]](#). Studi terbaru melaporkan bahwa SSAT diregulasi pada kanker paru-paru [\[62\]](#). 2-Deoksido-glukosa (2DG), analog glukosa, diubah menjadi 2-DG-P oleh heksokinase. 2-DG-P tidak dapat dimetabolisme tetapi dapat menekan heksokinase secara alosterik, yang merupakan enzim pembatas laju glikolisis [\[63\]](#). Karena memblokir glikolisis, pengaruh 2-DG dalam berbagai proses biologis. Ini dapat menghambat glikosilasi terkait-N, meningkatkan stres oksidatif, dan secara efisien menekan pertumbuhan dan invasi sel [\[64\]](#). Asam amino asam 2-aminoethanesulfonic, umumnya dikenal sebagai taurin, memiliki efek fisiologis yang luas dan dikonfirmasi sebagai bahan anti-cedera endogen [\[65\]](#). Obat ini dapat meningkatkan regulasi ekspresi N-acetyl galactosaminyl transferase 2, menurunkan regulasi ekspresi matriks metalloproteinase-2, dan menghambat potensi invasi dan metastasis [\[66\]](#). Penelitian sebelumnya telah mengusulkan bahwa perubahan kadar taurin dapat digunakan untuk memprediksi transformasi ganas dan pembentukan tumor payudara, kandung kemih, dan kolorektal [\[67–69\]](#).

Baru-baru ini, pada bulan Maret 2020, Chabon dkk. digunakan untuk mengintegrasikan fitur genom untuk deteksi dini kanker paru-paru non-invasif [\[70\]](#), yang awalnya mendemonstrasikan metode pembelajaran mesin dapat digunakan untuk mendeteksi kanker paru-paru. Berdasarkan fitur DNA bebas sel (cfDNA), para peneliti mengembangkan dan memvalidasi metode pembelajaran mesin secara prospektif

disebut 'kemungkinan kanker paru-paru dalam plasma' (Lung-CLiP), yang dapat membedakan pasien kanker paru stadium awal dengan pasien kontrol. Selama uji skrining, mereka mengamati sensitivitas 63% pasien kanker paru stadium I dengan spesifisitas 80%. Dibandingkan dengan model Naïve Bayes, Neural Network, dan SVM yang didasarkan pada biomarker metabolik sebagai fitur yang dikembangkan dalam penelitian kami, sensitivitas dan spesifisitasnya berbeda >85%. Meskipun model pembelajaran mesin kami kurang akurat dibandingkan LDCT, strategi ini berpotensi meningkatkan jumlah total pasien yang diskriminasi. Saat strategi ini berkembang menjadi uji klinis, data sampel yang berlimpah akan memungkinkan peningkatan kinerja dengan menggunakan algoritma pembelajaran mesin yang lebih progresif.

Di sisi lain, penelitian kami saat ini masih memiliki beberapa keterbatasan. Pertama, analisis kami didasarkan pada data dari satu institusi layanan kesehatan di wilayah geografis yang terbatas. Selain kanker paru-paru non-sel kecil, jumlah pasien dengan jenis kanker paru-paru lainnya tidak mencukupi. Oleh karena itu, diperlukan studi konfirmasi lebih lanjut di lembaga lain sebelum penerapannya. Kedua, data kami hanya mencakup tingkat metabolit. Informasi lebih lanjut tentang pasien tumor paru-paru dan individu sehat, seperti usia, riwayat merokok, diagnosis tumor yang bersamaan, dan riwayat kesehatan masa lalu, akan membantu untuk penelitian lebih lanjut. Kami mengusulkan bahwa integrasi biomarker metabolik plasma dengan skrining CT atau fitur tumor paru lainnya dapat lebih meningkatkan kinerja.

Bagaimanapun, kita perlu menemukan metode yang tepat untuk menerapkan strategi ini dalam praktik klinis, seperti menggabungkan sistem chip elektronik untuk melakukan pengujian plasma dan penerapan model di jalur perakitan. Salah satu penerapan potensial dari penelitian kami dapat digunakan sebagai skrining utama untuk beberapa pasien kanker paru-paru. Meskipun kandidat LDCT termasuk orang yang berisiko tinggi, pasien-pasien ini tidak diskriminasi karena kekhawatiran akan hasil positif palsu, akses terbatas, dan alasan terbatas lainnya. Kemudian pasien yang menunjukkan hasil tes positif akan dirujuk untuk pemeriksaan LDCT. Selain itu, dengan memodifikasi metode pembelajaran mesin dan menggabungkan fitur-fitur yang sesuai untuk jenis kanker lainnya, kami berharap dapat mengembangkan strategi yang menggabungkan metabolomik dan pembelajaran mesin untuk beragam diagnosis keganasan.

Kesimpulan

Metode interdisipliner perintis diusulkan dalam penelitian ini untuk mendeteksi biomarker diagnostik kanker paru-paru dini dengan menggabungkan metode metabolomik dan pembelajaran mesin. Biomarker metabolik menunjukkan kekuatan diagnostik yang signifikan untuk deteksi dini tumor paru-paru.

Deklarasi Kepentingan Bersaing

Para penulis menyatakan bahwa mereka tidak mempunyai kepentingan finansial atau hubungan pribadi yang saling bersaing yang dapat mempengaruhi pekerjaan yang dilaporkan dalam makalah ini.

Pernyataan kontribusi kepenulisan CRediT

Ying Xie: Metodologi, Investigasi, Penulisan - review & editing. **Wei-Yu Meng:** Penulisan - draf asli, Metodologi, Analisis formal. **Jalankan-Ze Li:** Penulisan - draf asli, Konseptualisasi, Investigasi. **Yu-Wei Wang:** Peringkat Lunak, Kurasi Data. **Xin Qian:** Investigasi, Sumber Daya. **Chang Chan:** Investigasi, Sumber Daya. **Zhi-Fang Yu:** Investigasi, Sumber Daya. **Penggemar Xing-Xing:** Sumber daya. **Hu-Dan Pan:** Sumber daya. **Chun Xie:** Sumber Daya, Administrasi proyek. **Qi-Biao Wu:** Sumber daya. **Pei-Yu Yan:** Sumber daya. **Liang Liu:** Sumber Daya, Pengawasan. **Yi-Jun Tang:** Sumber Daya, Pengawasan. **Xiao-Jun Yao:** Konseptualisasi, Penulisan - review & editing, Akuisisi pendanaan. **Mei-Fang Wang:** Konseptualisasi, Sumber Daya, Penulisan - review & editing. **Elaine Lai-Han Leung:** Konseptualisasi, Penulisan - review & editing, Akuisisi pendanaan.

Pendanaan

Pekerjaan ini didanai oleh The Science and Technology Development Fund, Macau SAR (Nomor file:0096/2018/A3,0003/2019/AKP, 0055/2020/A) dan Proyek Dana Penelitian Koperasi Cendekiawan NSFC di luar negeri dan Hong Kong dan Makau (Nomor proyek: 81828013).

Bahan pelengkap

Materi pelengkap terkait artikel ini dapat ditemukan, dalam versi online, di doi:10.1016/j.tranon.2020.100907.

Referensi

- [1] DS Ettinger, DE Wood, DL Aisner, W.Akerley, J.Bauman, LR Chirieac, TA D'Amico, MM DeCamp, TJ Dilling, M. Dobelbower, RC Doebele, R. Govindan, MA Gubens, M. Hennon, L. Horn, R. Komaki, RP Lackner, M. Lanuti, TA Leal, LJ Leisch, R. Lilienbaum, J. Lin, BW Loo Jr., R. Martins, GA Otterson, K. Reckamp, GJ Riely, SE Schild, TA Shapiro, J. Stevenson, SJ Swanson, K. Tauer, SC Yang, K. Gregory, M. Hughes, Kanker paru-paru bukan sel kecil, versi 5.2017, pedoman praktik klinis NCCN di bidang onkologi, J. Natl. Kompr. Jaringan Kanker. 15 (2017) 504–535.
- [2] Z. Chen, CM Fillmore, PS Hammerman, CF Kim, KK Wong, Kanker paru-paru non-sel kecil: serangkaian penyakit heterogen, Nat. Pendeta Kanker 14 (2014) 535–546.
- [3] X. Jin, SJ Yun, P. Jeong, IY Kim, WJ Kim, S. Park, Diagnosis kanker kandung kemih dan prediksi kelangsungan hidup melalui metabolisme urin, Oncotarget 5 (2014) 1635–1645.
- [4] H. Wang, J. Chen, Y. Feng, W. Zhou, J. Zhang, YU Yu, X. Wang, P. Zhang, (1) H analisis metabolisme ekstraseluler berbasis resonansi magnetik nuklir dari karsinoma skuamosa oral T8a113 yang resisten terhadap banyak obat sel, Onkol. Biarkan. 9 (2015) 2551–2559.
- [5] CW Lam, Hukum CY, Profil metabolik efusi pleura berbasis spektrometri massa yang tidak ditargetkan: asam lemak sebagai biomarker kanker baru untuk efusi pleura ganas, J. Proteome Res. 13 (2014) 4040–4046.
- [6] S. Bamji-Stocke, V. van Berkel, DM Miller, HB Frieboes, Tinjauan biomarker terkait metabolisme dalam diagnosis dan pengobatan kanker paru-paru, Metabolomics 14 (2018) 81.
- [7] J. Zhu, D. Djukovic, L. Deng, H. Gu, F. Himmati, EG Chiorean, D. Raftery, Deteksi kanker kolorektal menggunakan profil metabolik serum yang ditargetkan, J. Proteome Res. 13 (2014) 4120–4130.
- [8] W. Guan, M. Zhou, CY Hampton, BB Benigno, LD Walker, A. Gray, JF McDonald, FM Fernández, Deteksi kanker ovarium dari data kromatografi cair metabolik/ spektrometri massa dengan mesin vektor pendukung, BMC Bioinform. 10 (2009) 259.
- [9] K. Kim, P. Aronov, SO Zakharkin, D. Anderson, B. Perroud, IM Thompson, RH Weiss, Analisis metabolik urin untuk deteksi kanker ginjal dan penemuan biomarker, Mol. Proteom Sel. 8 (2009) 558–570.
- [10] S. Tiziani, V. Lopes, UL Günther, Diagnosis tahap awal kanker mulut menggunakan metabolisme berbasis 1H NMR, Neoplasia 11 (2009) 269–276 264p setelah 269.
- [11] S. Urayama, W. Zou, K. Brooks, V. Tolstikov, Profil metabolik plasma darah berbasis spektrometri massa yang komprehensif mengungkapkan pengklasifikasi diskriminatif yang kuat dari kanker pankreas, Rapid Commun. Spektrum Massa. 24 (2010) 613–620.
- [12] RD Nindrea, T. Aryandono, L. Lazuardi, I. Dwiprahasto, Keakuratan diagnostik sebagai algoritma pembelajaran mesin untuk penghitungan risiko kanker payudara: metaanalisis, Asian Pac. J. Kanker Sebelumnya. 19 (2018) 1747–1752.
- [13] J. Vamathevan, D. Clark, P. Czdrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, S. Zhao, Penerapan pembelajaran mesin dalam bidang obat-obatan penemuan dan pengembangan, Nat. Penemuan Narkoba Pendeta. 18 (2019) 463–477.
- [14] V. Dalal, J. Carmichael, A. Dhaliwal, M. Jain, S. Kaur, SK Batra, Radiomik dalam stratifikasi lesi kistik pankreas: pembelajaran mesin dalam aksi, Cancer Lett. 469 (2020) 228–237.
- [15] B. Zhang, X. He, F. Ouyang, D. Gu, Y. Dong, L. Zhang, X. Mo, W. Huang, J. Tian, S. Zhang, Pengklasifikasi pembelajaran mesin radiomik untuk biomarker prognostik karsinoma nasofaring stadium lanjut, Cancer Lett. 403 (2017) 21–27.
- [16] E. Mucaki, JZL Zhao, DJ Lizotte, PK Rogan, Memprediksi respons terhadap agen kemoterapi platin dengan pembelajaran mesin yang terinspirasi biokimia, Signal Transduct. Targetkan disana. 4 (2019) 1.
- [17] W. Xu, M. Xu, L. Wang, W. Zhou, R. Xiang, Y. Shi, Y. Zhang, Y. Piao, Analisis integratif metilasi DNA dan ekspresi gen mengidentifikasi biomarker diagnostik khusus kanker serviks, Transduksi Sinyal. Target. Ada. 4 (2019) 55.
- [18] S. Huang, J. Yang, S. Fong, Q. Zhao, Kecerdasan buatan dalam diagnosis dan prognosis kanker: peluang dan tantangan, Cancer Lett. 471 (2020) 61–71.
- [19] CM Lynch, B. Abdollahi, JD Fuqua, AR de Carlo, JA Bartholomai, RN Balgeman, VH van Berkel, HB Frieboes, Prediksi kelangsungan hidup pasien kanker paru-paru melalui teknik klasifikasi pembelajaran mesin yang diawasi, Int. J. Med. Memberitahukan. 108 (2017) 1–8.
- [20] S. Luo, J. Xu, Z. Jiang, L. Liu, Q. Wu, EL Leung, AP Leung, Metode pemfilteran kolaboratif berbasis kecerdasan buatan dengan pembelajaran ansambel untuk pengobatan kanker paru-paru yang dipersonalisasi tanpa pengurutan genetik, Pharmacol. Res. 160 (2020) 105037.
- [21] EI Emin, E. Emin, A. Papaloi, F. Willmott, S. Clarke, M. Sideris, Kecerdasan buatan dalam kebidanan dan ginekologi: apakah ini jalan ke depan? Di Vivo 33 (2019) 1547–1551.
- [22] WD Travis, E. Brambilla, AP Burke, A. Marx, AG Nicholson, Pengantar klasifikasi tumor paru-paru, pleura, timus, dan jantung Organisasi Kesehatan Dunia 2015, J. Thorac. Onkol. 10 (2015) 1240–1242.
- [23] R. Suguro, XC Pang, ZW Yuan, SY Chen, YZ Zhu, Y. Xie, Penerapan kombinasi silybin dan tangeretin melemahkan perkembangan steatohepatitis non-alkohol (NASH) pada tikus melalui modulasi metabolisme lipid, Pharmacol. Res. 151 (2020) 104519.
- [24] H. Pan, Y. Zheng, Z. Liu, Z. Yuan, R. Ren, H. Zhou, Y. Xie, L. Liu, Menguraikan mekanisme farmakologis Guan-jie-Kang dalam mengobati arthritis yang diinduksi bahan pembantu tikus menggunakan analisis omics, Depan. medis. 13 (2019) 564–574.
- [25] C. Chen, R. Xia, H. Chen, Y. He, TBtools: Perangkat untuk Ahli Biologi yang Mengintegrasikan Berbagai Alat Penanganan Data HTS dengan Antarmuka yang Ramah Pengguna, bioRxiv, (2018) 289660.
- [26] J. Damsar, T. Curk, A. Erjavec, C. Gorup, T. Hocevar, M. Milutinovic, dkk., Orange: kotak peralatan pengembangan data dengan Python, J. Mach. Pembelajaran. Res. 14 (2013) 2349–2353.
- [27] S. Huang, N. Cai, PP Pacheco, S. Narrandes, Y. Wang, W. Xu, Penerapan pembelajaran support vector machine (SVM) dalam genomik kanker, Genom Kanker. Proteom. 15 (2018) 41–51.
- [28] P. Shi, S. Ray, Q. Zhu, MA Kon, Pasangan dengan skor tertinggi untuk pemilihan fitur dalam pembelajaran mesin dan aplikasi pada prediksi hasil kanker, BMC Bioinform. 12 (2011) 375.
- [29] B. Ambale-Venkatesh, X. Yang, CO Wu, K. Liu, WG Hundley, R. McClelland, AS Gomes, AR Folsom, S. Shea, E. Guallar, DA Bluemke, JAC Lima, Prediksi kejadian kardiovaskular dengan pembelajaran mesin: studi multi-etnis tentang aterosklerosis, Circ. Res. 121 (2017) 1092–1101.
- [30] NA Zaidi, J. Cerquides, MJ Carman, GI Webb, Alleviating Naive Bayes asumsi independensi atribut dengan pembobotan atribut, J. Mach. Pembelajaran. Res. 14 (2013) 1947–1988.
- [31] AB Banu, P. Thirumalaikolundusubramanian, Perbandingan pengklasifikasi Bayes untuk klasifikasi kanker payudara, Asian Pac. J. Kanker Sebelumnya. 19 (2018) 2917–2920.
- [32] HH Rashidi, NK Tran, EV Betts, LP Howell, R. Green, Kecerdasan buatan dan pembelajaran mesin dalam patologi: lanskap metode yang diawasi saat ini, Acad. jalan. 6 (2019) 2374289519873088.
- [33] WA Cooper, S. O'Toole, M. Boyer, L. Horvath, A. Mahar, Apa yang baru dalam kanker paru-paru non-sel kecil bagi ahli patologi: pentingnya subtype yang akurat, mutasi EGFR dan penataan ulang ALK, Patologi 43 (2011) 103–115.
- [34] B. Liu, SJ Kim, KJ Cho, S. Oh, Pengembangan model pembelajaran mesin untuk diagnosis glaukoma, PLoS One 12 (2017) e0177726.
- [35] MS Ahmed, M. Shahjaman, MM Rana, MNH Mollah, Penguatan pengklasifikasi Naive Bayes dan penerapannya untuk analisis data ekspresi gen microarray, Biomed. Res. Int. 2017 (2017) 3020627.
- [36] H. Zhang, JX Ren, JX Ma, L. Ding, Pengembangan model prediksi in silico untuk toksisitas saluran kemih akibat bahan kimia dengan menggunakan pengklasifikasi Naive Bayes, Mol. Penyelam 23 (2019) 381–392.
- [37] T. Hubbard, D. Barker, E. Birney, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, R. Durbin, E. Eyras, J. Gilbert, M. Hammond, L. Huminecki, A. Kasprzyk, H. Lehtvaslaiho, P. Lijnzaad, C. Melsopp, E. Mongin, R. Pettett, M. Pocock, S. Potter, A. Rust, E. Schmidt, S. Searle, G. Slater, J. Smith, W. Spooner, A. Stabenau, J. Stalker, E. Stupka, A. Ureta-Vidal, I. Vastrik, M. Clamp, Proyek database genom Ensembl, Nucleic Acids Res. 30 (2002) 38–41.
- [38] F. Li, M. Yang, Y. Li, M. Zhang, W. Wang, D. Yuan, D. Zhang, Model prediksi stadium karsinoma sel ginjal sel jernih yang ditingkatkan berdasarkan set gen, BMC Bioinform. 21 (2020) 232.
- [39] V. Barroso-García, GC Gutiérrez-Tobal, L. Kheirandish-Gozal, D. Álvarez, F. Vaquerizo-Villar, P. Núñez, F. Del Campo, D. Gozal, R. Hornero, Kegunaan plot perulangan dari aliran udara rekaman untuk membantu diagnosis apnea tidur pediatrik, Comput. Metode Program Biomed. 183 (2020) 105083.
- [40] J. Ko, N. Bhagwat, T. Black, SS Yee, YJ Na, S. Fisher, J. Kim, EL Carpenter, BZ Stanger, D. Issadore, profil miRNA dari vesikel ekstraseluler yang diisolasi nanopori magnetik untuk diagnosis kanker pankreas, Cancer Res. 78 (2018) 3688–3697.
- [41] W. Zhao, J. Yang, Y. Sun, C. Li, W. Wu, L. Jin, Z. Yang, B. Ni, P. Gao, P. Wang, Y. Hua, M. Li, Pembelajaran mendalam 3D dari CT scan memprediksi invasi tumor adenokarsinoma paru subsentimeter, Cancer Res. 78 (2018) 6881–6889.
- [42] PF Pinsky, DS Gierada, W. Black, R. Munden, H. Nath, D. Aberle, E. Kazerooni, Kinerja RADS paru-paru dalam Uji Coba Skrining Paru-Paru Nasional: penilaian retrospektif, Ann. Magang. medis. 162 (2015) 485–491.
- [43] L. Li, J. Chen, X. Chen, J. Tang, H. Guo, X. Wang, J. Qian, G. Luo, F. He, X. Lu, Y. Ding, Y. Yang, W. Huang, G. Hou, X. Lin, Q. Ouyang, H. Li, R. Wang, F. Jiang, R. Pu, J. Lu, M. Jin, Y. Tan, FJ Gonzalez, G. Cao, M. Wu, H. Wen, T. Wu, L. Jin, L. Chen, H. Wang, Serum miRNA sebagai biomarker prediktif dan preventif untuk karsinoma hepatoseluler praklinis, Cancer Lett. 373 (2016) 234–240.
- [44] B. Poell, RJ van Haastert, F. Cerisoli, AS Bolijn, LM Timmer, B. Diosdado – Calvo, GA Meijer, AA van Puijenbroek, E. Berezikov, RQ Schaapveld, E. Cuppen, Skrining mikroRNA fungsional menggunakan mikroRNA manusia lentiviral yang komprehensif perpustakaan ekspresi, BMC Genom. 12 (2011) 546.
- [45] Q. Yang, P. Zhang, R. Wu, K. Lu, H. Zhou, Mengidentifikasi kombinasi penanda terbaik dalam CEA, CA125, CY211, NSE, dan SCC untuk skrining kanker paru-paru dengan menggabungkan kurva ROC dan analisis regresi logistik: adalah itu mungkin? Dis. Penanda 2018 (2018) 1–12.
- [46] S. Kitazono, Y. Fujiwara, K. Tsuta, H. Utsumi, S. Kanda, H. Horinouchi, H. Nokihara, N. Yamamoto, S. Sasada, S. Watanabe, H. Asamura, T. Tamura, Y. Ohe, Keandalan sampel biopsi kecil dibandingkan dengan spesimen yang direseksi untuk penentuan ekspresi ligan kematian terprogram 1 di paru-paru non-sel kecil kanker, Klinik. Kanker Paru-Paru 16 (2015) 385–390.
- [47] S. Riniker, Y. Wang, JL Jenkins, GA Landrum, Menggunakan informasi dari layar throughput tinggi historis untuk memprediksi senyawa aktif, J. Chem. Inf. Model. 54 (2014) 1880–1891.

- [48] J. Jeon, S. Nim, J. Teyra, A. Datti, J.L. Wrana, S.S. Sidhu, J. Moffat, P.M. Kim, Pendekatan sistematis untuk mengidentifikasi target obat kanker baru menggunakan pembelajaran mesin, desain inhibitor, dan skrining throughput tinggi, *Kedokteran Genom.* 6 (2014) 57.
- [49] B. Obrzut, M. Kusy, A. Semczuk, M. Obrzut, J. Kluska, Prediksi kelangsungan hidup keseluruhan 5 tahun pada pasien kanker serviks yang diobati dengan histerektomi radikal menggunakan metode kecerdasan komputasi, *BMC Cancer* 17 (2017) 840.
- [50] M. Kyrgiou, A. Pouliakis, J.G. Panayiotides, N. Margari, P. Bountris, G. Valasoulis, M. Paraskevaiddi, E. Bilirakis, M. Nasioutziki, A. Loufopoulos, M. Haritou, D.D. Koutsouris, P. Karakitsos, E. Paraskevaiddis, Manajemen personalisasi wanita dengan kelainan serviks menggunakan sistem penilaian pendukung keputusan klinis, *Gynecol. Onkol.* 141 (2016) 29–35.
- [51] Y. Huang, J. Zhu, W. Li, Z. Zhang, P. Xiong, H. Wang, J. Zhang, Panel mikroRNA serum digali dengan pembelajaran mesin sebagai biomarker potensial untuk mendeteksi kanker lambung, *Oncol. Rep.* 39 (2018) 1338–1346.
- [52] Y. He, M. Gao, H. Tang, Y. Cao, S. Liu, Y. Tao, Perantara metabolik dalam tumorigenesis dan perkembangan, *Int. J. Biol. Sains.* 15 (2019) 1187–1199.
- [53] M. Phang, Metabolisme prolin dalam regulasi sel dan biologi kanker: kemajuan dan hipotesis terkini, *Antioksidan. Sinyal Redoks.* 30 (2019) 635–649.
- [54] W. Liu, K. Glunde, Z.M. Bhujwalla, V. Raman, A. Sharma, J.M. Phang, Prolin oksidase meningkatkan kelangsungan hidup sel tumor di lingkungan mikro tumor hipoksia, *Cancer Res.* 72 (2012) 3677–3686.
- [55] W. Liu, C.N. Hancock, J.W. Fischer, M. Harman, J.M. Phang, Biosintesis prolin menambah pertumbuhan sel tumor dan glikolisis aerobik: keterlibatan nukleotida piridin, *Sci. Rep.* 5 (2015) 17206.
- [56] Y. Liu, C. Mao, M. Wang, N. Liu, L. Ouyang, S. Liu, H. Tang, Y. Cao, S. Liu, X. Wang, D. Xiao, C. Chen, Y. Shi, Q. Yan, Y. Tao, Perkembangan kanker dimediasi oleh katabolisme prolin pada kanker paru-paru non-sel kecil, *Onkogen* 39 (2020) 2358–2376.
- [57] L.R. Kolodziej, E.M. Paleolog, R.O. Williams, Metabolisme Kynurenine dalam kesehatan dan penyakit, *Asam Amino* 41 (2011) 1173–1183.
- [58] S.C. Chuang, A. Fanidi, P.M. Ueland, C. Relton, O. Midttun, S.E. Vollset, M.J. Gunter, M.J. Seckl, R.C. Travis, N. Wareham, A. Trichopoulou, P. Lagiou, D. Trichopoulos, P.H. Peeters, H.B. Bueno-de-Mesquita, H. Boeing, A. Wientzek, T. Kuehn, R. Kaaks, R. Tumino, C. Agnoli, D. Palli, A. Naccarati, E.A. Aicua, M.J. Sánchez, J.R. Quirós, M.D. Chirlaque, A. Agudo, M. Johansson, K. Grankvist, M.C. Boutron-Ruault, F. Clavel-Chapelon, G. Fagherazzi, E. Weiderpass, E. Riboli, P.J. Brennan, P. Vineis, M. Johansson, Mengedarkan biomarker triptofan dan jalur kynurenine serta risiko kanker paru-paru, *Cancer Epidemiol. biomark. Sebelumnya* 23 (2014) 461–468.
- [59] A.E. Pegg, Spermidin/spermine-N(1)-asetiltransferase: pengatur metabolisme utama, *Am. J. Fisiol. Endokrinol. Metab.* 294 (2008) E995–1010.
- [60] N. Babbar, A. Hacker, Y. Huang, R.A. Casero Jr., Faktor nekrosis tumor alfa menginduksi spermidine/spermin N1-asetiltransferase melalui faktor nuklir kappaB dalam sel kanker paru-paru non-sel kecil, *J. Biol. kimia.* 281 (2006) 24182–24192.
- [61] A.N. Kingsnorth, H.M. Wallace, Peningkatan poliamina monoasetilasi pada kanker payudara manusia, *Eur. J. Klinik Kanker. Onkol.* 21 (1985) 1057–1062.
- [62] S. Singhal, C. Rolfo, A.W. Maksymiuk, P.S. Tappia, D.S. Sitar, A. Russo, P.S. Akhtar, N. Khatun, P. Rahnuma, A. Rashiduzzaman, R. Ahmed Bux, G. Huang, B. Ramjiawan, Biopsi cair dalam skrining kanker paru-paru: kontribusi metabolomik. Hasil studi percontohan, *Cancer* 11 (2019) 1069.
- [63] M. Parniak, N. Kalant, Penggabungan glukosa menjadi glikogen dalam kultur primer hepatosit tikus, *Can. J. Biokimia. Biol. Sel.* 63 (1985) 333–340.
- [64] D. Zhang, J. Li, F. Wang, J. Hu, S. Wang, Y. Sun, 2-Deoxy-d-glucose menargetkan metabolisme glukosa dalam sel kanker sebagai terapi potensial, *Cancer Lett.* 355 (2014) 176–183.
- [65] X. Zhang, S. Tu, Y. Wang, B. Xu, F. Wan, Mekanisme apoptosis yang diinduksi taurin pada sel kanker usus besar manusia, *Acta Biochim. Biofisika. Dosa.* 46 (2014) 261–272.
- [66] P.M. Neary, P. Hallihan, J.H. Wang, R.W. Pfirrmann, D.J. Bouchier-Hayes, H.P. Redmond, Peran taurolidine yang berkembang dalam terapi kanker, *Ann. Bedah. Onkol.* 17 (2010) 1135–1143.
- [67] I.M. El Agouza, S.S. Eissa, M.M. El Houseini, D.E. El-Nashar, O.M. Abd El Hameed, Taurine: penanda tumor baru untuk meningkatkan deteksi kanker payudara di antara pasien wanita, *Angiogenesis* 14 (2011) 321–330.
- [68] S. Srivastava, R. Roy, S. Singh, P. Kumar, D. Dalela, S.N. Sankhwar, A. Goel, AA Sonkar, Taurine – kemungkinan biomarker sidik jari pada kanker kandung kemih invasif non-otot: studi percontohan dengan spektroskopi ¹H NMR, *Cancer Biomark.* 6 (2010) 11–20.
- [69] S. Tu, X.L. Zhang, H.F. Wan, Y.Q. Xia, Z.Q. Liu, X.H. Yang, F.S. Wan, Pengaruh taurin terhadap proliferasi sel dan apoptosis sel kanker paru-paru manusia A549, *Oncol. Biarkan.* 15 (2018) 5473–5480.
- [70] J. Chabon, E.G. Hamilton, D.M. Kurtz, M.S. Esfahani, M. Diehn, Mengintegrasikan fitur genom untuk deteksi dini kanker paru non-invasif, *Nature* 580 (2020) 245–251.