PENGKLASIFIKASIAN DEBITUR DENGAN MENGGUNAKAN ALGORITMA *GRAHAM SCAN* DALAM PENGAPLIKASIAN *CONVEX HULL*

AGUS EKA ARIESTA[1], G.K. GANDHIADI[2], NI KETUT TARI TASTRAWATI[3], I PUTU EKA NILA KENCANA[4]

[1,2,3,4]Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali
e-mail: [1]goesboe_bsm@yahoo.com, [2]gandhiadigk@yahoo.com, [3]taritastrawati@yahoo.com, [4]i_putu_enk@gmail.com

Abstract

Computational geometry is the mathematical science of computation by using the algorithm analysis to solve the problems of geometry. The problems of computational include polygon triangulations, convex hulls, Voronoi diagrams, and motion planning. Convex hull is the set of points that form a convex polygon that covers the entire set of points. The algorithms for determining the convex hull, among others, Graham Scan, Jarvis March, and Divide and Conquer. In the two-dimensional case, Graham Scan algorithm is highly efficient in the use of time complexity. This article discusses the quest convex hull of the data bank debtors, some of the data used to look at the classification accuracy of the convex hull formed. The coordinates of all the data found by using principal component analysis.After the data are analyzed, we get the accuracy of classification by 74%.

Keywords*: Computational geometry, convex hull, graham scan, principal component analysis*

1. Pendahuluan

Pada decade terakhir suatu disiplin ilmu yang dikenal dengan geometri komputasi telah muncul. Geometri komputasi merupakan suatu ilmu pada bidang matematika komputasi yang berhubungan dengan rancangan dan analisis algoritma untuk menyelesaikan masalah–masalah geometri [1]. Bidang yang mendasari geometri komputasi adalah matematika diskret dan algoritma geometrik.

Permasalahan geometri komputasi mencakup beberapa hal, yakni trigonometri poligon (*polygon triangulations*), *convex hulls*, *Voronoi diagrams*, *geometry searching*, dan *motion planning*. Salah satu permasalahan yang berkaitan dengan penentuan bentuk geometri komputasi dapat diselesaikan dengan menggunakan algoritma pencarian *convex hull*. Tujuan utama dari pencarian *convex hull* adalah menggambarkan perbedaan antarkelompok pada variabel-variabel kontinu [3]. Adapun algoritma untuk menentukan *convex hull* dari himpunan titik berhingga antara lain *Brute Force*, *Graham Scan*, *Jarvis March*,dan *Divide and Conquer*. Tiap algoritma memiliki kelebihan dan

[1] Mahasiswa Jurusan Matematika FMIPA Universitas Udayana
[2,3,4] Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

kekurangan masing-masing, namun dalam kasus berdimensi dua,algoritma *Graham Scan* merupakan algortima yang cukup efisien dalam kompleksitas waktu yang digunakan [2].

Dalam tulisan ini akan dibahas pencarian *convex hull* dari data debitur bank dengan menggunakan algoritma *Graham Scan*. Debitur ini dikelompokkan berdasarkan kemampuannya dalam melunasi kredit yakni kelompok lancar, diberi perhatian khusus, kurang lancar, diragukan, dan macet. Dari data yang tersedia, beberapa data akan digunakan untuk melihat seberapa besar ketepatan *convex hull* dalam mengklasifikasikan debitur berdasarkan kemampuan dalam melunasi kreditnya.

2. Metode Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data besarnya pinjaman debitur, jangka waktu pinjaman, penghasilan per bulan debitur, pendidikan terakhir debitur, dan kemampuan debitur dalam melunasi kreditnya yang berjumlah 200 data.

Adapun teknik analisis data pada penelitian ini dilakukan dengan langkah-langkah berikut:

1. Melakukan transformasi data
   Transformasi dilakukan pada peubah bebas yang berskala ordinal, dengan tujuan untuk mengubah data yang berskala ordinal tersebut menjadi berskala interval dengan *Method of Successive Interval* (MSI). Pada penelitianini, transformasi data dilakukan hanya pada variable pendidikan terakhir. Adapun langkah-langkahnya sebagai berikut:
   a. Dihitung frekuensi untuk setiap kategori.
   b. Berdasarkan frekuensi setiap kategori dihitung proporsinya.
   c. Dari proporsi yang diperoleh, dihitung proporsi kumulatif untuk setiap kategori.
   d. Ditentukan nilai batas Z untuk setiap kategori.
   e. Dihitung *scale value* untuk setiap kategori melalui persamaan:
   $$scale_{value} = \frac{density\ at\ lower\ limit - density\ at\ upper\ limit}{area\ under\ upper\ limit - area\ under\ lower\ limit}$$
   f. Dihitung *score* untuk setiap kategori melalui persamaan:
   $$score = scale_{value} + \left|scale_{valu\,e_{min}}\right| + 1$$
2. Mencari titik koordinat (x, y) masing-masing data dengan menggunakan analisis komponen utama dengan mereduksikan variabel X_1, X_2, X_3, X_5. Adapun langkah-langkahnya sebagai berikut:
   a. Menentukan matrik sragam$[\Sigma]$ yaitu matriks $X^{'}X$.
   b. Mencari nilai eigen λ dari matriks ragam$[\Sigma]$.
   c. Mencari dua vector eigen yang bersesuaian dengan dua nilai eigen terbesar.

d. Menentukan *score* komponen vector cirri pada langkah *c*. *Score* komponen pertama sebagai absis (x) dan *score* komponen kedua sebagai ordinat (y).

3. Data tersebut akan dikelompokkan menjadi dua kelompok dengan menggunakan teknik pengelompokan *stratified random sampling*. Kelompok data pertama sebanyak 150 digunakan untuk membangun *convex hull* (data analisis) dan kelompok data kedua sebanyak 50 berikutnya (data validasi) digunakan untuk memvalidasi ketepatan klasifikasi.
4. Dua kelompok yang telah terbentuk pada langkah 3 akan dikelompokkan lagi masing-masing menjadi lima subkelompok yang baru berdasarkan kemampuan debitur dalam melunasi kreditnya, yakni kelompok lancar, diberi perhatian khusus, kurang lancar, diragukan, dan macet.
5. Setiap subkelompok akan dibentuk pemetaan berdimensi dua dengan menggunakan titik koordinat (x, y) yang diperoleh dari langkah 2.
6. Pembentukan *convex hull*.

   Adapun algoritma pembentukan *convex hull* dengan algoritma *Graham Scan* dijelaskan sebagai berikut:

   a) Setiap data pada masing-masing kelompok diurutkan berdasar kanordinat y terkecil. Jika terdapat lebih dari satu ordinat y terkecil maka ordinat y terkecil dengan absis x terkecil dipilih menjadi titik awal pengurutan (p_0).
   b) Pada masing-masing kelompok dicari besar sudut berlawanan arah jarum jam yang terbentuk dari titik awal pengurutan(p_0) dan sumbu mendatar yang melalui titik awal pengurutan terhadap titik-titik lainnya.
   c) Data diurutkan berdasarkan sudut yang diperoleh pada masing-masing kelompok dari terkecil sampai terbesar (*ascending*).
   d) Pada masing-masing kelompok dengan data yang telah terurut dilakukan proses scan. Dipilih tiga buah titik terkecil (terurut) untuk diuji nilai hasil kali silangnya. Jika hasil kali silang tersebut bernilai positif maka proses scan dilanjutkan ketitik berikutnya, sedangkan jika nilai hasil kali silang tiga titik bernilai negative atau nol maka titik terkecil kedua dari tiga titik yang dipilih dihapus. Proses scan dilakukan sampai ketitik terakhir padamasing-masing kelompok. Sehingga diperoleh titik-titik luar yang merupakan titik pembentuk *convex hull* pada masing-masing kelompok.
7. Menentukan kelompok masing-masing individu data validasi berdasarkan *convex hull* yang terbentuk. Jika terdapat data yang berada pada wilayah *overlapping* (tumpang tindih) dan di luar *convex hull* yang terbentuk, maka data-data tersebut akan diklasifikasikan kembali berdasarkan criteria berikut:

a) Pengklasifikasian data yang berada pada wilayah *overlapping* akan ditentukan dengan mencari jarak *Euclid* terpendek antara data terhadap *centroid-centroid* (titik tengah) wilayah yang bersangkutan.
b) Pengklasifikasian data yang berada di luar *convex hull* yang terbentuk akan ditentukan denganjarak *Euclid* terpendek antara data terhadap *centroid* kelima *convex hull* yang terbentuk

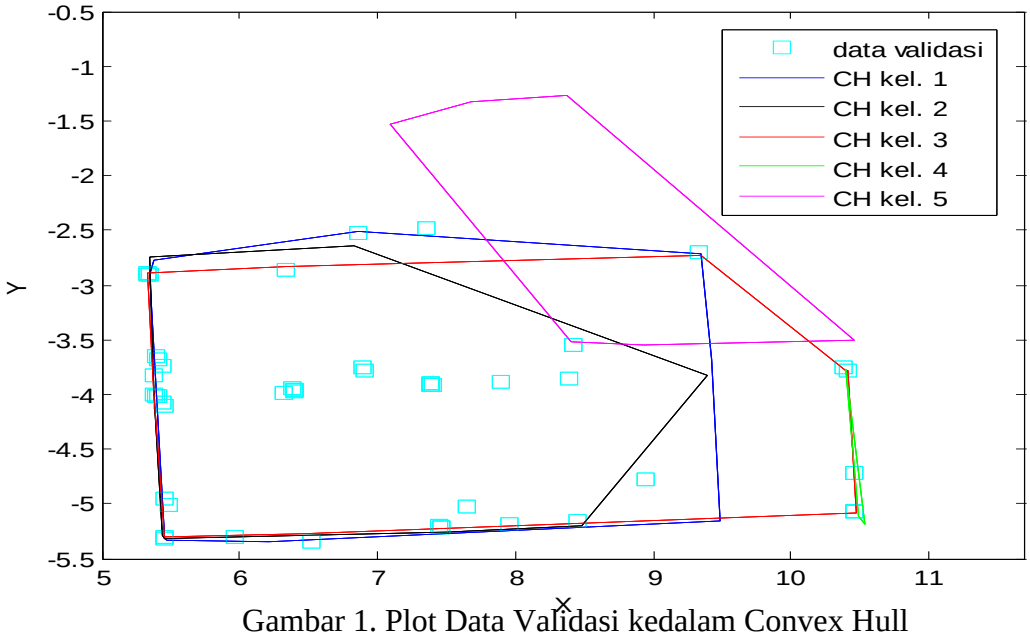
8. Mengukur ketepatan pengklasifikasian dengan menggunakan data validasi.
9. Lakukan pengulangan langkah ke-3 sampai ke-9 sebanyak sembilan kali, untuk mencari pengelompokan mana yang mempunyai ketepatan klasifikasi terbesar.

3. Hasil dan Pembahasan

Langkah awal yang dilakukan adalah transformasi variabel yang berskala ordinal menjadi skala interval dengan menggunakan *Method of Successive Interval* (MSI). Variabel yang ditransformasi adalah variable pendidikan terakhir. Setelah diubah menjadi skala interval, langkah selanjutnya adalah mencari titik koordinat(x, y) yang ditentukan oleh dua *score* component terbesar pada Analisis Komponen Utama, *score* komponen terbesar pertama sebagai absis(x) dan *score* komponen kedua sebagai ordinat (y).

Setelah didapatkan titik koordinat (x, y), selanjutnya 200 data tersebut dibagi menjadi dua yakni 150 sebagai pembentuk *convex hull* dan 50 sebagai data validasi. Pembagian kelompok ini menggunakan teknik *stratified random sampling*. Data yang berjumlah 200 akan dibagi menjadi lima kelompok yaitu, 90 data kelompok lancar, 46 data kelompok diberiperhatian khusus, 51 data kelompok kurang lancar, 4 data kelompok diragukan, dan 9 data kelompok macet. Selanjutnya pada masing-masing kelompok dilakukan pengacakan secara *random*, lalu dipilih data yang akan dijadikan data analisis sampel dan data validasi sampel.

Selanjutnya dibentuk pemetaan berdimensi dua dari data analisis yang telah diolah. Kemudian membentuk *convex hull* masing-masing kelompok dengan algoritma *Graham Scan*. Setelah *convex hull* terbentuk, data validasi dimasukkan kedalam *convex hull* untuk menentukan ketepatan klasifikasi data tersebut.



Gambar 1. Plot Data Validasi kedalam Convex Hull

Keterangan :

= data validasi

= wilayah yang membentuk kelompok lancar

= wilayah yang membentuk kelompok diberi perhatian khusus

= wilayah yang membentuk kelompok kurang lancar

= wilayah yang membentuk kelompok diragukan

= wilayah yang membentuk kelompok macet

Perhatikan Gambar 1, terdapat beberapa data yang berada pada wilayah *overlapping* (tumpang tindih) dan diluar *convex hull* yang terbentuk. Data yang berada pada wilayah yang *overlapping* maupun diluar *convex hull* akan diklasifikasikan kembali. Adapun pengklasifikasian untuk data-data tersebut ditentukan berdasarkan kriteria berikut:

1. Pengklasifikasian data yang berada pada wilayah *overlapping* akan ditentukan dengan mencari jarak *Euclid* terpendek antara data terhadap *centroid-centroid* (titik tengah) wilayah yang bersangkutan.
2. Pengklasifikasian data yang berada di luar *convex hull* yang terbentuk akan ditentukan dengan mencari jarak *Euclid* terpendek antara data terhadap *centroid* dari kelima *convex hull* yang terbentuk.

Jarak *Euclid* dari data terhadap titik *centroid* yang bersesuaian(JEC) dapat dicari dengan rumus $JEC_j = \sqrt{(x_i - x_{cj})^2 + (y_i - y_{cj})^2}$dengan,

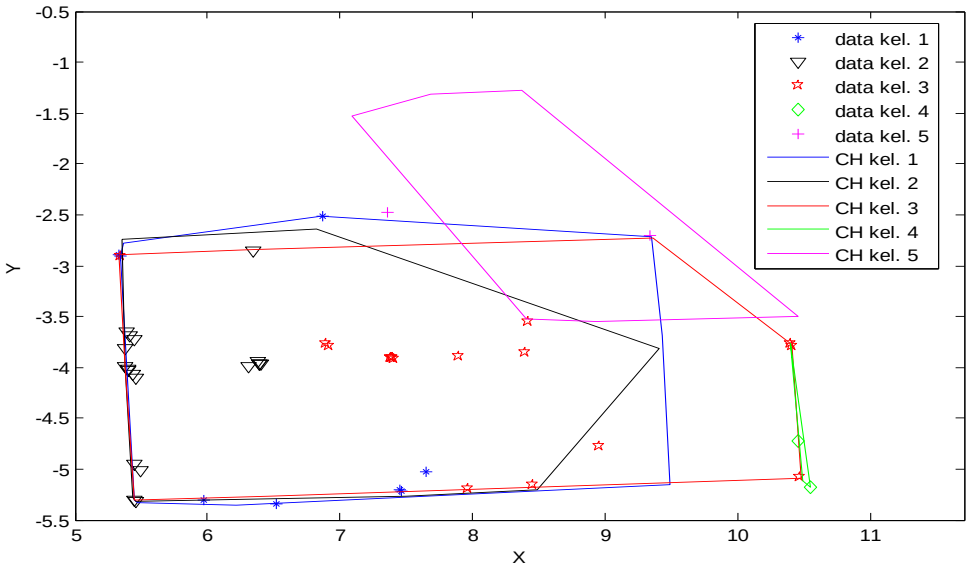
x_i = komponen xd ata ke-i $(i = 1, \dots, 150)$

y_i = komponen y data ke-i

x_{cj} = jarak *centroid* komponen x pada kelompok data ke-$j(i = 1, \dots, 5)$

y_{cj} = jarak *centroid* komponen y pada kelompok data ke-j

Hasil dari data yang telah diklasifikasikan kembali disajikan pada Gambar 2.



Gambar 2. Hasil Akhir Pengklasifikasian kedalam Convex Hull

Berdasarkan Gambar 2 diperoleh hasil ketepatan pengklasifikasian sebesar 74%. Hasil ketepatan ini adalah hasil yang terbesar setelah melakukan proses mencari ketepatan pengklasifikasian sebanyak sepuluh kali.Adapun ringkasan dari sepuluh pengulangan disajikan pada Tabel 1.

Tabel 1. Hasil Sepuluh Pengulangan Proses Mencari Ketepatan Pengklasifikasian

Pengulangan	Data Berada pada		Ketepatan Klasifikasi
	Wilayah yang *overlapping*	Di Luar *Convex Hull*	
1	40	5	60%
2	31	16	58%
3	36	1	74%
4	39	9	54%
5	37	4	68%
6	37	9	60%
7	38	9	58%
8	34	6	68%
9	35	9	66%
10	34	5	70%

Dari Tabel 1 dapat dilihat bahwa ketepatan klasifikasi dipengaruhi oleh banyaknya data yang berada pada wilayah yang *overlapping* dan di luar *convex hull* yang terbentuk.

4. Kesimpulan

Berdasarkan hasil yang diperoleh, dapat diambil kesimpulan bahwa dari setiap pengulangan diperoleh hasil ketepatan pengklasifikasian yang berbeda-beda, dan ketepatan terbesar adalah 74%. Hasil ketepatan pengklasifikasian ini dipengaruhi oleh banyaknya data validasi yang berada pada wilayah *overlapping* dan data yang berada diluar *convex hull* yang terbentuk. Semakin banyak data yang berada pada wilayah *overlapping* dan yang berada diluar *convex hull*, maka ketepatan pengklasifikasiannya cenderung lebih kecil.

Daftar Pustaka

[1] Johnsonbaugh, Richard. 2002. *Discrete Mathematics*. Fourth edition. Penerjemah: Didiek Djunaedi. Jakarta :PT Prehallindo.

[2] O'Rourke, Joseph. 1997. *Computational Geometry in C*. Second edition. Cambrige University Press.

[3] Vidmar,G. M. P. 2004. *Augmented Convex Hull Plots: Rationale, Implementation in R and Beomedical Applications.* Slovenia: Institute of Biomedical Informatics Faculty of Medicine University of Ljubljana.