



*Building
Future
Leaders*

PROPOSAL PROGRAM KREATIVITAS MAHASISWA
“KLASIFIKASI KANKER PAYUDARA MENGGUNAKAN TEKNIK
***NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE*”**

BIDANG KEGIATAN
PKM PENELITIAN

Diusulkan oleh:

Zulfikar Akbar	3145161915	2016
Muhammad Insan Khamil	3145161580	2016
Agus Setiawan	3145161716	2016

UNIVERSITAS NEGERI JAKARTA
JAKARTA
2017

PENGESAHAN USULAN PKM PENELITIAN EKSAKTA

- | | |
|--|--|
| 1. Judul Kegiatan | : KLASIFIKASI KANKER PAYUDARA
MENGUNAKAN TEKNIK NAIVE BAYES DAN
SUPPORT VECTOR MACHINE |
| 2. Bidang Kegiatan | : PKMPE - Kesehatan |
| 3. Ketua Pelaksana Kegiatan | |
| a. Nama Lengkap | : ZULFIKAR AKBAR |
| b. NIM | : 3145161915 |
| c. Program Studi | : S1 Ilmu Komputer |
| d. Perguruan Tinggi | : Universitas Negeri Jakarta |
| e. Alamat Rumah dan No Telp./HP | : Kp. Tegal Rotan Rt 002/007 No. 25, Sawah Baru, Ciputat,
Tangerang Selatan, Banten 15413, telp. -, hp.
087808085169 |
| f. Alamat email | : zulfikar.78.akbar@gmail.com |
| 4. Anggota Pelaksana Kegiatan/Penulis | : 2 orang |
| 5. Dosen Pendamping | |
| a. Nama Lengkap dan Gelar | : MUHAMMAD EKA SURYANA S.Kom, M.Kom |
| b. NIDN | : 0323128503 |
| c. Alamat Rumah dan No Telp./HP | : Komplek IKPN Blok E No.20, Bintaro, Jakarta Selatan,
telp. -, hp. 085891166321 |
| 6. Biaya Kegiatan Total | |
| a. Dikti | : Rp 7,963,000 |
| b. Sumber lain (sebutkan...) | : Rp 0; Sumber lain: - |
| 7. Jangka Waktu Pelaksanaan | : 3 bulan. |

Jakarta, 30 - 11 - 2017

Menyetujui

Wakil Dekan Bidang Kemahasiswaan, Alumni, dan Ketua Pelaksana Kegiatan,
Kerja Sama Universitas Negeri Jakarta,




(Dr. Iwan Sugibartono, M. Si.)
NIP/NIK. 19791010 200801 1 108




(ZULFIKAR AKBAR)
NIM. 3145161915

Wakil Rektor Bidang Kemahasiswaan dan Alumni
Universitas Negeri Jakarta,



(Prof. Dr. Achmad Sofyan Hanif, M. Pd.)
NIP/NIK. 19630917 198903 1 002

Dosen Pendamping,



(MUHAMMAD EKA SURYANA S.Kom, M.Kom)
NIDN. 0323128503

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PENGESAHAN	ii
DAFTAR ISI	iii
DAFTAR TABEL DAN GAMBAR	iv
BAB 1. PENDAHULUAN	1
BAB 2. TINJAUAN PUSTAKA	2
BAB 3. METODE PENELITIAN	7
BAB 4. BIAYA DAN JADWAL KEGIATAN	8
4. 1 Anggaran Biaya	8
4. 2 Jadwal Kegiatan	8
DAFTAR PUSTAKA	9
LAMPIRAN-LAMPIRAN	11
Lampiran 1. Biodata Ketua, Anggota, dan Dosen Pendamping	11
Lampiran 2. Justifikasi Anggaran Kegiatan	16
Lampiran 3. Susunan Organisasi Tim Peneliti dan Pembagian Tugas	18
Lampiran 4. Surat Pernyataan Ketua Peneliti	19

DAFTAR TABEL

Tabel 2. 1 Deskripsi data	2
Tabel 2. 2 Informasi atribut	2
Tabel 2. 3 <i>Confusion matrix</i>	5
Tabel 4. 1 Anggaran Biaya	8
Tabel 4. 2 Jadwal Kegiatan	8

DAFTAR GAMBAR

Gambar 2. 1 SVM berusaha mencari <i>hyperplane</i> terbaik	4
Gambar 2. 2 SVM mendapatkan <i>hyperplane</i> terbaik	4
Gambar 3. 1 Prosedur penelitian	7

BAB 1. PENDAHULUAN

Berdasarkan data WHO tahun 2013, insiden kanker meningkat dari 12,7 juta kasus tahun 2008 menjadi 14,1 juta kasus tahun 2012, dengan jumlah kematian meningkat dari 7,6 juta orang tahun 2008 menjadi 8,2 juta pada tahun 2012. Kanker menjadi penyebab kematian nomor 2 di dunia sebesar 13% setelah penyakit kardiovaskular (Kemenkes RI, 2014a). Prevalensi kanker di Indonesia sebesar 1,4 per 1000 penduduk. Dilihat dari karakteristik jenis kelamin penderita kanker di Indonesia, perempuan sebesar 2,2 per 1.000 penduduk dan laki-laki sebesar 0,6 per 1.000 penduduk. Berdasarkan estimasi Globocan, *International Agency for Research on Cancer* (IARC) tahun 2012 dalam Kemenkes RI (2014a), kanker payudara adalah jenis kanker dengan presentase kasus baru tertinggi (43,3%) dan presentase kematian tertinggi (12,9%) pada perempuan di dunia.

Pemeriksaan payudara sendiri (SADARI) dilakukan untuk menurunkan angka mortalitas kanker payudara dengan penemuan kanker payudara sedini mungkin dan pengobatan saat ukuran masih kecil sebelum kanker tersebut bermetastasis (Blamey, 2000). Penemuan kanker payudara sedini mungkin yang didiagnosis dan diobati secara benar akan menambah harapan hidup penderita kanker payudara. Angka harapan hidup selama 10 tahun untuk penemuan kanker pada stadium I sebesar 70%-80%, stadium II 43%, stadium III kurang dari 11,2%, dan stadium IV 0% (Soetjipto, 2001).

Pada penelitian oleh Leena Vig (2014) menyajikan sebuah analisis menggunakan *Random Forest Classifier*, *Artificial Neural Network* (ANN), *Naive Bayes* dan *Support Vector Machine* (SVM). Hasil menunjukkan bahwa ANN, *Random Forest* dan SVM mampu untuk menghasilkan model dengan akurasi, sensitivitas, dan spesifisitas tinggi sedangkan *Naive Bayes* buruk.

Penelitian yang dilakukan menggunakan *Wisconsin Breast Cancer dataset* yang dapat diunduh dari *website UCI Machine Learning Repository*. Dua teknik klasifikasi *Naive Bayes* dan *Support Vector Machine* (SVM) bisa digunakan sebagai metode pada *machine learning*. Metode *Naive Bayes* yaitu metode klasifikasi probabilistik berdasarkan Teorema Bayes. *Support Vector Machine* yaitu menggunakan *hyperplane* untuk memisahkan beberapa contoh yang berbeda ke dalam kelas masing-masing. Data dipartisi dengan perbandingan 90% untuk pelatihan dan 10% untuk pengujian. Untuk mengukur kinerja, maka digunakan *10-fold cross validation* dan *confusion matrix*. Untuk melakukan penelitian ini media yang digunakan ialah bahasa pemrograman Python yang mempunyai paket teknik klasifikasi menggunakan *machine learning*. Hasil penelitian ini diharapkan membantu pencegahan dini supaya angka kematian akibat kanker payudara di Indonesia berkurang.

BAB 2. TINJAUAN PUSTAKA

A. Data

Data medis tentang kanker payudara (*breast cancer*) dari *University of Wisconsin Hospital* merupakan data yang diperoleh dari 699 pasien yang telah diklasifikasikan ke dalam 2 (dua) jenis kanker, yaitu kanker jinak (*benign*) dan kanker ganas (*malignant*). *Database* tersebut berupa 11 atribut yang mewakili data untuk setiap pasien, yang terdiri dari 9 (sembilan) karakteristik dari sel payudara dan 2 (dua) atribut lainnya, yaitu nomor identitas dari setiap pasien dan kelas yang sesuai dengan jenis kanker payudara (jinak atau ganas). Variabel yang terdapat pada data diperoleh dari hasil FNA (*Fine-Needle Aspirate*) *Biopsy Cytology* yang merupakan salah satu jenis biopsi yang dilakukan dengan cara mengambil sampel sel pada benjolan atau massa payudara yang dicurigai sebagai tumor atau kanker dengan menggunakan jarum halus untuk selanjutnya dilakukan pemeriksaan di laboratorium (Wolberg & Mangasarian, 1990)

Tabel 2. 1 Deskripsi data

No	Data	Deskripsi
1	Banyak data	699
2	Banyak atribut	10
3	Atribut 2 sampai 10	Data
4	Kelas	<i>Benign</i> <i>Malignant</i>
5	Distribusi kelas	<i>Benign</i> : (458) (65,5%) <i>Malignant</i> : (241) (34,5%)

Tabel 2. 2 Informasi atribut

No	Atribut	Domain
1	<i>Sample code number</i>	Nomor Id
2	<i>Clump Thickness</i>	1 - 10
3	<i>Uniformity of Cell Size</i>	1 - 10
4	<i>Uniformity of Cell Shape</i>	1 - 10
5	<i>Marginal Adhesion</i>	1 - 10
6	<i>Single Epithelial Cell Size</i>	1 - 10
7	<i>Bare Nuclei</i>	1 - 10
8	<i>Bland Chromatin</i>	1 - 10

9	<i>Normal Nucleoli</i>	1 - 10
10	<i>Mitoses</i>	1 - 10
11	<i>Class</i>	2 untuk <i>Benign</i> 4 untuk <i>Malignant</i>

Berikut merupakan penjelasan terkait dengan parameter-parameter pada *Wisconsin Breast Cancer Database* (Salama et al, 2012: 37-38):

1. *Clump Thickness*

Pada atribut ini, sel *benign* cenderung dikelompokkan dalam monolayer (lapisan tunggal), sementara sel-sel kanker sering dikelompokkan dalam multilayer (lapisan yang lebih dari satu).

2. *Uniformity of Cell Size and Shape*

Pada atribut ini, sel kanker cenderung bervariasi dalam ukuran dan bentuk. Parameter ini menjadi alasan untuk menentukan jenis sel, kanker atau tidak.

3. *Marginal Adhesion*

Pada atribut ini, sel normal cenderung untuk tetap bersama-sama, sedangkan pada sel-sel kanker cenderung menyebar. Jadi hilangnya adhesi adalah tanda keganasan.

4. *Single Epithelial Cell Size*

Atribut ini terkait dengan keseragaman yang telah dijelaskan sebelumnya. Sel-sel epitel yang membesar secara signifikan kemungkinan dapat menjadi sel kanker.

5. *Bare Nuclei*

Bare Nuclei adalah istilah yang digunakan untuk inti sel yang tidak dikelilingi oleh sitoplasma. *Bare Nuclei* biasanya terlihat pada tumor jinak.

6. *Bland Chromatin*

Atribut ini menggambarkan keseragaman 'tekstur' dari inti sel yang terlihat pada sel benign. Sedangkan pada sel-sel kanker, kromatin cenderung bertekstur kasar.

7. *Normal Nucleoli*

Normal Nucleoli adalah struktur kecil yang terlihat di inti sel (nukleus). Dalam sel-sel normal, nukleolus biasanya sangat kecil jika terlihat. Sedangkan pada sel kanker nukleolus menjadi lebih menonjol, dan terkadang jumlahnya lebih dari jumlah normal.

8. *Mitoses*

Mitosis adalah proses dimana sel membagi dan membelah diri. Kelas kanker dapat ditentukan dengan menghitung jumlah mitosis.

B. *Naïve Bayes*

Naïve Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas (Tina R. Patil dan S.S. Sherekar, 2013). Definisi lain mengatakan *Naïve Bayes* merupakan

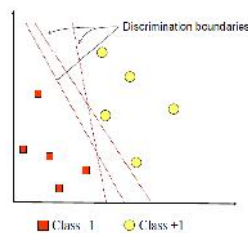
pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya (Bustami, 2013).

Keuntungan penggunaan *Naïve Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naïve Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan (Shadab Adam Pattekari dan Asma Parveen, 2012).

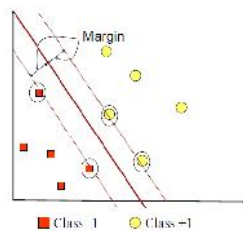
C. Support Vector Machine (SVM)

Dalam SVM, untuk memisahkan data terhadap kelasnya, SVM akan membangun sebuah *hyperplane* (bidang pemisah). Sebuah *hyperplane* (bidang pemisah) yang baik, bukan hanya *hyperplane* yang bisa digunakan untuk memisahkan data, akan tetapi *hyperplane* yang baik adalah *hyperplane* yang memiliki batasan (*margin*) yang paling besar. Pencarian bidang pemisah terbaik inilah yang menjadi inti dari *Support Vector Machine*. Akan tetapi, dalam proses pencarian *hyperplane* tersebut, akan muncul permasalahan baru yaitu sebuah formula yang sangat sulit untuk dipecahkan, yang disebut dengan permasalahan *Quadratic Programming* (Munawarah, Raudlatul. 2016).

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* (Tsuda, 2000) terbaik yang berfungsi sebagai pemisah dua buah kelas pada *input space*. *Pattern* yang merupakan anggota dari dua buah kelas: +1 dan -1 dan berbagi alternatif garis pemisah (*discrimination boundaries*). *Margin* adalah jarak antara *hyperplane* tersebut dengan *pattern* terdekat dari masing-masing kelas. *Pattern* yang paling dekat ini disebut sebagai *support vector*. Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pembelajaran pada SVM (Christianini, Nello dan John S. Taylor. 2000).



Gambar 2. 1 SVM berusaha mencari *hyperplane* terbaik
(Sumber: Nugroho, A. S., Arief Budi Witarto, dan Dwi Handoko, 2003)



Gambar 2. 2 SVM mendapatkan *hyperplane* terbaik
(Sumber: Nugroho, A. S., Arief Budi Witarto, dan Dwi Handoko, 2003)

E. *Confusion Matrix*

Confusion matrix merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi binari. Pada dasarnya *confusion matrix* mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya (E. Prasetyo, 2012).

Pada klasifikasi binari, data masukan dikelompokkan ke dalam salah satu dari dua kelas. Jenis klasifikasi ini merupakan bentuk klasifikasi yang paling sederhana dan banyak digunakan. Contoh penggunaannya antara lain dalam sistem yang melakukan deteksi orang atau bukan. Pada penelitian ini digunakan klasifikasi binari dengan alasan data yang digunakan nantinya akan diklasifikasikan menjadi dua kelas yaitu *benign* dan *malignant* yang diinterpretasikan dengan angka 2 (dua) dan 4 (empat).

Pada pengukuran kinerja menggunakan *confusion matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Nilai *True Negative* (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan *False Positive* (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, *True Positive* (TP) merupakan data positif yang terdeteksi benar. *False Negative* (FN) merupakan kebalikan dari *True Positive*, sehingga data positif, namun terdeteksi sebagai data negatif.

Berdasarkan nilai *True Negative* (TN), *False Positive* (FP), *False Negative* (FN), dan *True Positive* (TP) dapat diperoleh nilai akurasi, presisi (*Positif Predictive Value*), *recall* (*True Positive rate* atau *Sensitivity*), *False Positive rate*, *True Negative rate* (*Specificity*), dan *False Negative rate*. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Nilai akurasi dapat diperoleh dengan persamaan (1). Nilai presisi menggambarkan jumlah data kategori positif yang diklasifikasikan secara benar dibagi dengan total data yang diklasifikasi positif. Presisi dapat diperoleh dengan persamaan (2). Nilai *recall* menunjukkan berapa persen data kategori positif yang terklasifikasikan dengan benar oleh sistem. Nilai *recall* diperoleh dengan persamaan (3). Nilai *False Positive rate* menunjukkan banyak data dalam kategori negatif yang salah diklasifikasikan menjadi positif, bisa dilihat pada persamaan (4). Nilai *True Negative rate* menunjukkan banyaknya data yang dikategorikan negatif dan diklasifikasi dengan benar, dapat dilihat pada persamaan (5). Nilai *False Negative rate* menunjukkan banyak data dalam kategori positif yang salah diklasifikasikan menjadi negatif, ditunjukkan oleh persamaan (6).

Tabel 2. 3 *Confusion matrix*

Keadaan sebenarnya	Prediksi	
	<i>True</i>	<i>False</i>
<i>True</i>	TP	FN
<i>False</i>	FP	TN

$$A = \frac{T + T}{T + T + F + F} 100\% \dots\dots\dots(1)$$

$$P = \frac{T}{T + F} 100\% \dots\dots\dots(2)$$

$$R = \frac{T}{T + F} 100\% \dots\dots\dots(3)$$

$$F r = \frac{F}{F + T} 100\% \dots\dots\dots(4)$$

$$T r = \frac{T}{F + T} 100\% \dots\dots\dots(5)$$

$$F r = \frac{F}{F + T} 100\% \dots\dots\dots(6)$$

Pada penelitian ini, hanya nilai sensitivitas (*recall*) dan spesifisitas (*True Negative rate*) yang digunakan untuk mengukur kinerja metode klasifikasi *Naïve Bayes* dan *Support Vector Machine*. Sebagai tambahan, pada klasifikasi binari *recall* dari kelas yang positif disebut juga *sensitivity* dan *recall* dari kelas negatif disebut *specificity* (*scikit-learn Python user guide 0.20*).

F. *K-fold cross validation*

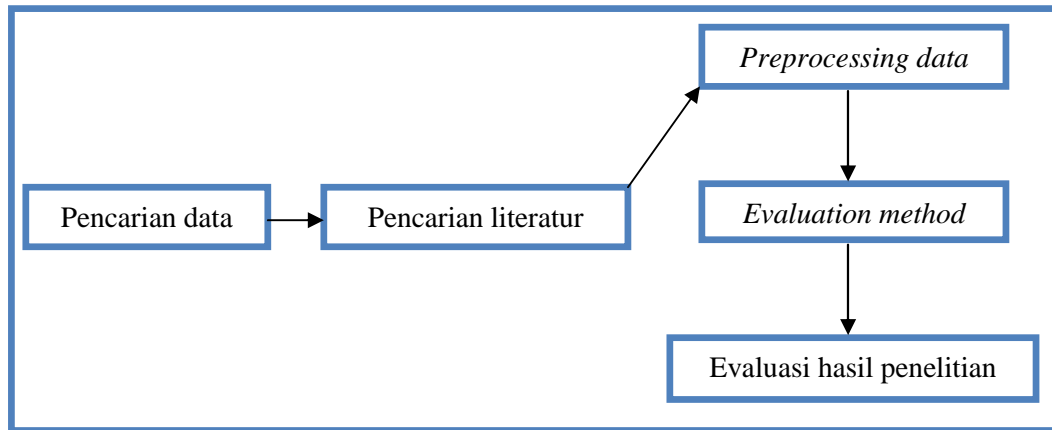
K-Fold Cross Validation yang merupakan salah satu metode yang digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak (Pandie, 2012).

K-Fold Cross Validation mengulang k-kali untuk membagi sebuah himpunan contoh secara acak menjadi k subset yang saling bebas, setiap ulangan disisakan satu subset untuk pengujian dan subset lainnya untuk pelatihan (Fu, 1994).

Untuk memperkirakan tingkat kesalahan yang terjadi dapat digunakan $k = 5$ atau 10, sebab data *training* pada setiap fold cukup berbeda dengan data *training* yang asli. Secara keseluruhan, 5 atau 10-fold *cross validation* sama-sama direkomendasikan dan disepakati bersama (Hastie, et al, 2008). Pada penelitian ini digunakan 10-fold *cross validation* sebagai metode untuk evaluasi teknik klasifikasi data kanker payudara.

BAB 3. METODE PENELITIAN

3. 1 Prosedur Penelitian



Gambar 3. 1 Prosedur Penelitian

3. 1. 1 Pencarian data

Data yang digunakan adalah data kanker payudara dari *University of Wisconsin Hospital*. Data ini bersumber dari *website UCI Machine Learning Respository* yang dapat diunduh.

3. 1. 2 Pencarian literatur

Tahap ini mencari jurnal-jurnal terdahulu terkait dengan metode-metode yang dipakai pada penelitian ini. Pada tahap ini juga digunakan *scikit-learn Python user guide 0.20.0* dan *Pandas Python Toolkit Documentation 0.17.0* untuk membantu proses penelitian berupa kodingan *machine learning* menggunakan bahasa pemrograman Python.

3. 1. 3 *Preprocessing data*

Preprocessing data diawali dengan membaca data. *Preprocessing data* berkaitan dengan pengecekan data untuk mengetahui ada atau tidaknya data yang hilang (*missing value*). *Missing value* sebaiknya dibersihkan agar memudahkan proses selanjutnya.

3. 1. 4 *Evaluation method*

Pada tahap ini ada beberapa proses yaitu:

- 1) Menyeleksi data.
- 2) Menguji tingkat akurasi dari teknik *Naïve Bayes* dan *Support Vector Machine (SVM)* menggunakan *k-fold cross validation* dengan nilai $k = 10$.
- 3) Menguji tingkat sensitivitas dan spesifisitas data yang dimodelkan menggunakan teknik klasifikasi *Naïve Bayes* dan *Support Vector Machine (SVM)* menggunakan *confusion matrix*.

3. 1. 5 Evaluasi hasil penelitian

Pada tahap ini evaluasi dilakukan untuk mengetahui teknik klasifikasi terbaik berdasarkan perbandingan tingkat akurasi, sensitivitas, dan spesifisitas.

BAB 4. BIAYA DAN JADWAL KEGIATAN

4. 1 Anggaran Biaya

Penggunaan anggaran yang dibutuhkan dalam penelitian ini adalah Rp7.963.000,00 dengan rincian sebagai berikut:

Tabel 4. 1 Anggaran Biaya

No.	Jenis Pengeluaran	Biaya (Rp)
1.	Peralatan Penunjang	
	a. Laptop Asus A46CB	5.500.000,00
	b. Modem Bolt	500.000,00
	c. Paket Internet Modem Bolt (16 GB)	600.000,00
	d. Flashdisk	100.000,00
	Jumlah (Rp)	6.700.000,00
2.	Bahan Habis Pakai	
	a. Kertas A4 (1 rim)	50.000,00
	b. Refil tinta printer HP (<i>black</i>)	63.000,00
	c. Refil tinta printer HP (<i>color</i>)	180.000,00
	d. Jilid	5.000,00
	e. Meterai	65.000,00
	Jumlah (Rp)	363.000,00
3.	Perjalanan	
	Biaya transportasi individu	450.000,00
	Jumlah (Rp)	450.000,00
4.	Biaya Lain-Lain	
	Kuota handphone	450.000,00
	Jumlah (Rp)	450.000,00
	Total Jumlah (Rp)	7.963.000,00

4. 2 Jadwal Kegiatan

Tabel 4. 2 Jadwal Kegiatan

No.	Jenis Kegiatan	Bulan		
		1	2	3
1.	Mencari literatur yang sesuai dengan penelitian			
2.	Melakukan penelitian			
3.	Menyusun laporan			

DAFTAR PUSTAKA

- Wisconsin Breast Cancer dataset.*
[https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+\(original\)](https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original))
- O. L. Mangasarian and W. H. Wolberg. 1990. "Cancer diagnosis via linear programming", SIAM News, Volume 23, Number 5, September 1990, pp 1 & 18.
- William H. Wolberg and O.L. Mangasarian. 1990. "Multisurface method of pattern separation for medical diagnosis applied to breast cytology", Proceedings of the National Academy of Sciences, U.S.A., Volume 87, December 1990, pp 9193 9196.
- Tina R. Patil, S.S. Sherekar. 2013. *Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification*, International Journal of Computer Science and Applications, Vol. 6, No. 2, April 2013.
- Shadab Adam Pattekari, Asma Parveen. 2012, *Prediction System for Heart Disease Using Naive Bayes*, International Journal of Advanced Computer and Mathematical Sciences, ISSN 2230-9624, Vol. 3, Issue 3, 2012
- Vijayakumar S, Wu S. 1999. "Sequential Support Vector Classifiers and Regression", Proc. International Conference on Soft Computing (SOCO'99), Genoa, Italy, pp.610 619, 1999
- Vapnik V.N. 1999. "The Nature of Statistical Learning Theory", 2nd edition, Springer Verlag, New York Berlin Heidelberg, 1999
- Gouda I. Salama, Abdelhalim and Magdy Abd-elghany Zeid. 2012. "Breast Cancer Diagnosis On Three Different Datasets Using Multi-Classifiers", International Journal of Computer and Information Technology (2277 0764) Volume 01– Issue 01, September 2012.
- Leena Vig. 2014. "Comparative Analysis of Different Classifiers for the Wisconsin Breast Cancer Dataset", Open Access Library Journal, Volume 1 | e660, 2014.
- Bustami. 2013. *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi*, TECHSI: Jurnal Penelitian Teknik Informatika.
- Munawarah, Raudlatul. 2016. *Implementasi Metode Support Vector Machine untuk Mendiagnosa Penyakit Hepatitis*. Program S-1 Ilmu Komputer, Universitas Lambung Mangkurat: Banjarbaru.
- Blamey RW, Wilson ARM, Patnick J. 2000. ABC of breast diseases: Screening for breast cancer. British Medical Journal. 2000;16(321) September: 689-93.
- Soetjipto. 2001. Deteksi Dini dan Pengobatan Kanker Payudara. Makalah disampaikan dalam ceramah sehari di RS Kanker Dharmais, Jakarta, 6 Maret 2001
- Nugroho A.S., Arief Budi Witarto, Dwi Handoko. 2003. "Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika–", <http://ilmukomputer.com>
- Pandie, Emerensye S. Y. 2012. Implementasi Algoritma Data mining K-Nearest Neighbour (KNN) Dalam Pengambilan Keputusan Pengajuan Kredit. Jurusan Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana: Kupang

- Hastie Trevor, Tibshirani Robert, dan Jerome Friedman. 2008. The Elements of Statistical Learning Data Mining, Inference, and Prediction. California: Springer.
- Tsuda K. 2000. "Overview of Support Vector Machine", Journal of IEICE, Vol.83, No.6, 2000, pp.460-466 (in Japanese)
- E. Prasetyo. 2012. *Data Mining: Konsep dan Aplikasi menggunakan Matlab*, 1 ed. Yogyakarta: Andi Offset

LAMPIRAN-LAMPIRAN

Lampiran 1. Biodata Ketua, Anggota, dan Dosen Pendamping

A. Identitas Diri

1	Nama Lengkap	Zulfikar Akbar
2	Jenis Kelamin	Laki-laki
3	Program Studi	Ilmu Komputer
4	NIM	3145161915
5	Tempat dan Tanggal Lahir	Jakarta, 30 Maret 1998
6	E-mail	zulfikar.78.akbar@gmail.com
7	Nomor Telepon/HP	087808085169

B. Riwayat Pendidikan

	SD	SMP	SMA
Nama Institusi	SD N Ciputat IX	SMP N 178 Jakarta	SMA N 86 Jakarta
Jurusan	-	-	IPA
Tahun Masuk-Lulus	2004-2010	2010-2013	2013-2016

C. Pemakalah Seminar Ilmiah (*Oral Presentation*)

No	Nama Pertemuan Ilmiah/Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	-	-	-

D. Penghargaan dalam 5 tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)

No	Jenis Penghargaan	Institusi Pemberi Penghargaan	Tahun
1	-	-	-

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi. Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah Proposal Program Kreativitas Mahasiswa bidang Penelitian

Jakarta, 23 November 2017

Pengusul


Zulfikar Akbar

A. Identitas Diri

1	Nama Lengkap	Muhammad Insan Khamil
2	Jenis Kelamin	Laki-laki
3	Program Studi	Ilmu Komputer
4	NIM	3145161580
5	Tempat dan Tanggal Lahir	Boyolali, 21 Februari 1998
6	E-mail	Khamil_insan0@yahoo.com
7	Nomor Telepon/HP	081293402756

B. Riwayat Pendidikan

	SD	SMP	SMA
Nama Institusi	MI N 15 Bintaro	SMP N 178 Jakarta	SMA N 87 Jakarta
Jurusan	-	-	IPA
Tahun Masuk-Lulus	2004-2010	2010-2013	2013-2016

C. Pemakalah Seminar Ilmiah (*Oral Presentation*)

No	Nama Pertemuan Ilmiah/Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	-	-	-

D. Penghargaan dalam 5 tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)

No	Jenis Penghargaan	Institusi Pemberi Penghargaan	Tahun
1	-	-	-

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi. Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah Proposal Program Kreativitas Mahasiswa bidang Penelitian

Jakarta, 23 November 2017

Pengusul,



Muhammad Insan Khamil

A. Identitas Diri

1	Nama Lengkap	Agus Setiawan
2	Jenis Kelamin	Laki-laki
3	Program Studi	Ilmu Komputer
4	NIM	3145161716
5	Tempat dan Tanggal Lahir	Jakarta, 23 Agustus 1998
6	E-mail	agus2429e@gmail.com
7	Nomor Telepon/HP	0895374596745

B. Riwayat Pendidikan

	SD	SMP	SMA
Nama Institusi	SD N Lenteng Agung 07 Pagi	SMP N 242 Jakarta	SMA N 64 Jakarta
Jurusan	-	-	IPA
Tahun Masuk-Lulus	2004-2010	2010-2013	2013-2016

C. Pemakalah Seminar Ilmiah (*Oral Presentation*)

No	Nama Pertemuan Ilmiah/Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	-	-	-

D. Penghargaan dalam 5 tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)

No	Jenis Penghargaan	Institusi Pemberi Penghargaan	Tahun
1	-	-	-

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi. Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah Proposal Program Kreativitas Mahasiswa bidang Penelitian

Jakarta, 23 November 2017

Pengusul,



Agus Setiawan

A. Identitas Diri

1	Nama Lengkap	Muhammad Eka Suryana, M. Kom.
2	Jenis Kelamin	Laki-laki
3	Program Studi	Ilmu Komputer
4	NIDN	0323128503
5	Tempat dan Tanggal Lahir	Jakarta, 23 Desember 1985
6	E-mail	eka-suryana@unj.ac.id
7	Nomor Telepon/HP	085891166321

B. Riwayat Pendidikan

	S-1	S-2
Nama Institusi	Universitas Indonesia	Universitas Indonesia
Jurusan	Ilmu Komputer	Ilmu Komputer
Tahun Masuk-Lulus	2003-2009	2010-2012

C. Pemakalah Seminar Ilmiah (*Oral Presentation*)

No	Nama Pertemuan Ilmiah/Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	<i>International Conference on Advanced Computer Science and Information Systems (ICACSIS)</i>	<i>Analysis of Multi Codebook GLVQ versus standard GLVQ in discriminating Sleep Stages (2nd Author)</i>	Jakarta, Desember 2012
2	<i>International Conference on Advanced Computer Science and Information Systems (ICACSIS)</i>	<i>Modified Fuzzy-Neuro Generalized Learning Vector Quantization for Early Detection of Arrhythmias (2nd Author)</i>	Jakarta, Desember 2012
3	<i>International Conference on System, Man, and Cybernetics (SMC)</i>	<i>FNGLVQ FPGA Design for Sleep Stages Classification based on Electrocardiogram Signal (1st Author)</i>	Seoul, 2012
4	Konferensi Nasional Penelitian Matematika dan Pembelajarannya	Digitalisasi Proses Absensi Berbasis Sistem Informasi	Surakarta, 2016

		di Lingkungan UNJ: Tinjauan dari sisi arsitektur dan implementasi	
--	--	--	--

D. Penghargaan dalam 5 tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)

No	Jenis Penghargaan	Institusi Pemberi Penghargaan	Tahun
1	-	-	-

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi. Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah Proposal Program Kreativitas Mahasiswa bidang Penelitian

Jakarta, 23 November 2017

Pendamping,



Muhammad Eka Suryana, M. Kom.

Lampiran 2. Justifikasi Anggaran Kegiatan

1. Peralatan penunjang

Material	Justifikasi Pemakaian	Volume	Harga satuan (Rp)	Jumlah biaya (Rp)
Laptop Asus A46CB	Untuk mencari literatur dan membuat laporan	1	5.500.000,00	5.500.000,00
Modem Bolt	Untuk mencari literatur	1	500.000,00	500.000,00
Paket Internet Modem Bolt (16 GB)	Untuk mencari literatur	3	200.000,00	600.000,00
Flashdisk	Menyimpan data laporan	1	100.000,00	100.000,00
SUB TOTAL (Rp)				6.700.000,00

2. Bahan habis pakai

Material	Justifikasi Pemakaian	Volume	Harga satuan (Rp)	Jumlah biaya (Rp)
Kertas A4 (1 rim)	Untuk Mencetak laporan	1	50.000,00	50.000,00
refil tinta printer HP (<i>black</i>)	Mencetak laporan	3	31.000,00	63.000,00
refil tinta printer HP (<i>color</i>)	Mencetak laporan	3	60.000,00	180.000,00
Jilid	Menjilid laporan	1	5.000,00	5.000,00
Meterai	Sebagai bukti pernyataan	10	6.500,00	65.000,00
SUB TOTAL (Rp)				363.000,00

3. Perjalanan

Material	Justifikasi Pemakaian	Volume	Harga satuan (Rp)	Jumlah biaya (Rp)
Biaya transportasi individu	Bersama-sama menyusun laporan	3	150.000,00	450.000,00
SUB TOTAL (Rp)				450.000,00

4. Biaya lain-lain

Material	Justifikasi Pemakaian	Volume	Harga satuan (Rp)	Jumlah biaya (Rp)
Kuota handphone	Komunikasi antaranggota	3	150.000,00	450.000,00
SUB TOTAL (Rp)				450.000,00
TOTAL (Keseluruhan)				7.963.000,00

Lampiran 3. Susunan Organisasi Tim Peneliti dan Pembagian Tugas

No .	Nama/NRM	Program Studi	Bidang Ilmu	Alokasi Waktu (Jam/Minggu)	Uraian Tugas
1	Zulfikar Akbar/ 3145161915	S-1 Ilmu Komputer	Komputer	8 jam/minggu	Ketua pelaksana, Koordinator , Meneliti
2	Muhammad Insan Khamil/ 3145161580	S-1 Ilmu Komputer	Komputer	8 jam/minggu	Meneliti, Menyusun laporan
3	Agus Setiawan/ 3145161716	S-1 Ilmu Komputer	Komputer	8 jam/minggu	Meneliti, Menyusun laporan

Lampiran 4. Surat Pernyataan Ketua Peneliti



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
UNIVERSITAS NEGERI JAKARTA

Kampus Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta 13220
Telepon/Faximile : Rektor : (021) 4893854, PR I : 4895130, PR II : 4893918, PR III : 4892926, PR IV : 4893982
BAUK : 4750930, BAAK : 4759081, BAPSI : 4752180
Bagian UHTP : Telepon. 4893726, Bagian Keuangan : 4892414, Bagian Kepegawaian : 4890536, Bagian HUMAS : 4898486
Laman : www.unj.ac.id

SURAT PERNYATAAN KETUA PELAKSANA

Yang bertanda tangan dibawah ini :

Nama : Zulfikar Akbar

NRM : 3145161915

Prodi : Ilmu Komputer

Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa proposal PKM-P saya dengan judul **Klasifikasi Kanker Payudara Menggunakan Teknik Naïve Bayes dan Support Vector Machine** yang diusulkan untuk tahun ajaran 2017 Bersifat original dan belum pernah dibiayai oleh lembaga atau sumber dana lainnya.

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku dan mengembalikan seluruh biaya PKM-P yang sudah diterima ke kas Negara.

Demikian Pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Jakarta, 20 November 2017

Mengetahui,

Wakil Dekan Bidang Kemahasiswaan,
Alumni dan Kerjasama



Dr. Iwan Sugihartono, M.Si
NIP. 197910102008011018

Ketua Pelaksana



Zulfikar Akbar
3145161915