

Telekomunikazio Teknologiaren Ingeniaritzako Gradua

Gradu amaierako lana

FIBRILAZIO AURIKULAR PAROXISTIKOA AURRESATEKO METODOA

Ikaslea: Eguskiza Garcia, Itziar

Zuzendaria: Irusta Zarandona, Unai

2017-2018

AURKIBIDEA

LABURPENA	3
RESUMEN	4
SUMMARY	5
TAULEN ZERRENDA	6
IRUDIEN ZERRENDA.....	7
LABURDUREN ZERRENDA.....	8
1. SARRERA.....	9
2. TESTUINGURUA.....	10
3. HELBURUAK.....	13
3.1. HELBURU NAGUSIA	13
3.2. HELBURU PARTZIALAK.....	13
3.2.1. DATU-BASEA DEFINITU.....	13
3.2.2. SEINALEAK BISTARATZEKO TRESNAK GARATU.....	13
3.2.3. ALGORITMOA GARATU	14
3.2.4. EMAITZEN AZTERKETA ESTATISTIKOA BURUTU	14
4. ONURAK	15
4.1. ONURA SOZIALA	15
4.2. ONURA TEKNIKOAK	15
4.3. ONURA EKONOMIKOAK	15
5. ARTEAREN EGOERA.....	17
5.1. EKG SEINALEA.....	17
5.2. HEART RATE VARIABILITY	19
5.3. SAILKATZAILEAK	21
5.4. FA AURREATEKO METODO EZBERDINAK.....	22
6. METODOLOGIA.....	23
6.1. DATU-BASEA SORTZEAREN PROZESUA	23
6.1.1. OINARRIZKO DATU-BASEAREN AZALPENA.....	23
6.1.2. PROIEKTUAN ERABILIKO DEN DATU-BASEAREN GARAPENA	24
6.2. BISTARATZEKO TRESNAK.....	26
6.3. PARAMETRIZAZIOA	28
6.4. PARAMETROEN SAILKATZE GAITASUNA	29
6.4.1. PARAMETRO INDIBIDUALEN AZTERKETA.....	29

6.4.2.	PARAMETROEN ARTEKO ERLAZIOA.....	31
6.4.3.	SAILKATZAILEAK	32
6.5.	EMAITZEN LABURPENA	35
7.	PLANGINTZA.....	37
7.1.	LAN-TALDEA	37
7.2.	PROIEKTUA GARATZEKO GAUZATUTAKO FASEAK	37
7.2.1.	PROIEKTUAREN KUDEAKETA.....	37
7.2.2.	DOKUMENTAZIOAREN IRAKURKETA ETA MATLAB IKASTAROA.....	38
7.2.3.	BIHURKETEN GARAKETA	38
7.2.4.	DATU-BASEA GARATZEA	39
7.2.5.	PARAMETROEN KALKULUEN BURUTZEA.....	39
7.2.6.	ALGORITMOEN GARAPENA ETA EMAITZEN KALKULUA.....	40
7.2.7.	DOKUMENTAZIOA	41
7.3.	MUGARRIAK	41
8.	AURREKONTUA.....	43
8.1.	AURREKONTUAREN KONTU-SAILAK.....	43
8.1.1.	BARNE ORDUAK.....	43
8.1.2.	AMORTIZAZIOAK	43
8.1.3.	AZPIKONTRATAZIOAK.....	44
8.1.4.	GASTUAK	44
8.1.5.	AURREKONTU OSOAREN LABURPENA	44
9.	ARRISKUEN ANALISIA	45
9.1.	KOSTUEN ARRISKUAK (A)	45
9.2.	EGUTEGI ARAZOAK (B)	45
9.3.	DATUEN GALERA (C).....	46
9.4.	ARRISKU TEKNIKOAK (D)	46
9.5.	ARRISKU TEKNOLOGIKOAK (E)	46
9.6.	ARRISKU ANALISIAREN LABURPENA	46
10.	ONDORIOAK	48
11.	ERREFERENTZIAK.....	49
	ERANSKINAK.....	52
I.	ERANSKINA: ERABILTZAILEAREN GIDA.....	52
II.	EMAITZA ZABALAGOAK.....	59

LABURPENA

Bihotzeko arritmiak bihotzaren funtzio elektrikoan gertatzen diren irregulartasunak dira. Bihotzeko erritmoaren aldaketa hauek, erritmoa bizkortzeagatik ala gutxiagotzeagatik izan daitezke, eta ez dira nahitaez irregularrak izan behar. Hala ere, normalean arritmiek erritmo irregularra sortzen dute, bihotzaren berezko taupadan edota bihotzaren sistema elektrikoan anomaliak agertzen direnean.

Zenbait arritmia mota aurki daitezke; horien artean, aurikuletan agertzen dena bereiz daiteke. Kasu honetan, aurikuletan ematen diren taupadak desorekatuak direnean fibrilazio aurikularra (FA) agertu dela esaten da.

FA bihotz-arritmia ohikoena dela esan ohi da, baita osasun sisteman kosturik handiena eragiten duena ere. Izan ere, azkeneko hamarkadan FAren prebalentzia handituz joan da, eta okerrera jarraituko du gizartearen zahartzearen ondorioz.

FA ez da arritmien artean larriena, arritmia mota honek ez baitu berez gizakiaren bizitza arriskuan ipintzen. Hala ere, oso kaltegarria da bizitzarako. Izan ere, arritmia honek pertsona batengan bihotz-gutxiegitasunagatik edo enbolia batengatik morbiditate arrisku altua izatea eragin ditzake.

Hau guztia dela eta, FA auresateko metodoak izatea oso garrantzitsua dela ondoriozta dezakegu. Hori izango da, beraz, proiektu honetan zehar garatuko dena. Horretarako, zenbait pazienteren informazioaz baliatuko da, geroago azalduko den moduan, informazio hori elektrokardiograma deituriko erregistroetan oinarrituta egongo da. Beraz, erregistro horiek abiapuntutzat edukita, eta informazio jakin hori prozesatu eta landu ondoren, FA auresateko zenbait algoritmo garatuko dira. Azkenik, garatutako algoritmo horiekin lortutako emaitzak aztertuko dira, metodo hauek fidagarriak diren ala ez ondorioztatzeko.

RESUMEN

Se le llaman arritmias a las irregularidades que suceden en el sistema de conducción cardíaco. Estas variaciones que se dan en el ritmo cardíaco pueden ser porque el corazón acelere o disminuya su frecuencia cardíaca, y no tienen que ser necesariamente irregulares. Sin embargo, normalmente las arritmias crean ritmos irregulares cuando suceden anomalías en el marcapasos fisiológico del corazón o en el sistema eléctrico del corazón.

Se diferencian varias arritmias entre las cuales se encuentra la que sucede en las aurículas. En este caso, cuando la conducción de los latidos que se dan en estas cavidades es desordenada, se padece fibrilación auricular (FA).

La FA es la arritmia más común, además de ser también la que más gastos ocasiona en el sistema de la salud. De hecho, en la última década la prevalencia de dicha arritmia ha ido aumentando, y se prevé que irá a peor por consecuencia del envejecimiento de la población.

La FA no se encuentra entre las arritmias más graves, ya que esta arritmia no pone por si sola en riesgo la vida del ser humano. Sin embargo, es muy perjudicial para la salud ya que las personas que padecen esta enfermedad podrían tener un alto riesgo de morbilidad en caso de sufrir una insuficiencia cardíaca o una embolia.

Por todo esto, podemos deducir que es muy importante encontrar métodos que predican la FA. Por lo cual, eso va a ser lo que se va a desarrollar en este proyecto. Para ello, aprovecharemos de la información de ciertos pacientes la cual, como se explicará más adelante, se basa en registros llamados electrocardiogramas. Por consiguiente, teniendo estos registros como punto de partida, y después de haberlos procesado y trabajado, se desarrollarán diferentes algoritmos para predecir la FA. Por último, se analizarán los resultados conseguidos a partir de estos algoritmos, los cuales nos ayudaran a concluir si estos métodos son fiables o no.

SUMMARY

Arrhythmias are known as irregularities that happen in the cardiac conduction system. These heart rate variations can occur when the heart rate is accelerated or decreased, and they do not necessarily have to be irregular. However, arrhythmias usually create irregular rhythms when abnormalities occur in the physiological pacemaker of the heart or in the electrical system of the heart.

There are several types of arrhythmias, some of the most typical ones are those occurring in the atria. A patient suffers atrial fibrillation (AF) when the beats that occur in those cavities are disorganized.

AF is the most common arrhythmia, as well as being the one that causes most expenses in the health system. In fact, in the last decade the prevalence of this arrhythmia has been increasing, and it is expected that it will get worse due to the aging of the population.

AF is not one of the most serious arrhythmias, since it doesn't put human lives at risk. However, it is very detrimental for the patient's health, since people suffering from this disease could have a high risk of morbidity in case of suffering heart failure or an embolism.

Thus, we can conclude that it is very important to find methods that predict AF. Therefore, that's what is going to be developed in this project. To obtain this goal, we will take some information of certain patients which, as it will be explained later, is based on records called electrocardiograms. Therefore, after processing this information, different algorithms will be developed to predict AF. Finally, the results obtained from these algorithms will be analyzed, which will help us to conclude if these methods are reliable or not.

TAULEN ZERRENDA

Taula 1. Proiektuan parte hartzen duen lan-taldearen deskribapena	37
Taula 2. Mugarriak	41
Taula 3. Barne orduen kostuak	43
Taula 4. Amortizazioen kostuak	43
Taula 5. Gastuen laburpen taula	44
Taula 6. Guztizko kostu aitortpena	44
Taula 7. Parametro hoberenak eta bere AUC balioa	60
Taula 8. Korrelatutako parametro bikoteak eta haien korrelazio koefizientea	62

IRUDIEN ZERRENDA

Irudia 1. Goiko grafikan EKG normala ikus daiteke non P uhina argi eta garbi bereiz dezakegun. Azpiko grafikan, aldiz, FA daukan paziente baten EKG-aren adibidea ikusten da.	10
Irudia 2. FA-ren eraginez sortzen diren mugimendu desorekatuak	11
Irudia 3. EKG seinalea	18
Irudia 4. EKG seinaletik lortutako RR tarteak	18
Irudia 5. RR tarte jakin batekin lortutako Poincaré Plot	21
Irudia 6. Holter proba gauzatzeko gailua	22
Irudia 7. Jatorrizko fitxategiek jarraitutako bihurketa prozedura	24
Irudia 8. Erregistro jakin baten EKG seinalearen bi kanalen irudikapena 10 segundoko denbora tartean	25
Irudia 9. Datuak bistartzeko GUI-aren diseinua	27
Irudia 10. Analisisirako GUI-aren diseinua	28
Irudia 11. Kasu bakoitza sailkatzeko irizpidea	30
Irudia 12. Denborazko domeinuko parametro baten ROC kurba eta bere AUC balioa	31
Irudia 13. Algoritmorako erabilitako lau parametroren arteko korrelazio matrizea	32
Irudia 14. Algoritmoa garatzeko jarraitu den eskema	33
Irudia 15. SVM ereduaren adibidea	34
Irudia 16. Erabaki-zuhaitzen teknika	35
Irudia 17. 7 parametro onenen arteko erlazio azterketa	36
Irudia 18. Gantt diagrama	42
Irudia 19. Proiektuaren arriskuekin erlazionatutako probabilitate-eragin matrizea	47
Irudia 20. Proiektua garatzeko jarraitu den eskema	52
Irudia 21. afDetection hasierako leihoa	53
Irudia 22. Menu irristakorrak	53
Irudia 23. EKG seinalearen kanal ezberdinen aukeraketa	54
Irudia 24. Bost segundoko tarte batean antzemandako bi kanalen adibidea	54
Irudia 25. RR edo HR tarteak bistartzeko aukera	55
Irudia 26. Seinaleak ehun segundoko iraupenarekin bistartzearen adibidea	55
Irudia 27. R uhinen aldiuneak adierazteko aukera	55
Irudia 28. Paziente jakin baten lehenengo erregistroa R uhinen aldiuneekin	55
Irudia 29. FAP duten pazienteen erregistroen arteko desberdintasuna	56
Irudia 30. analysisMethods hasierako leihoa	57
Irudia 31. Paziente bakoitzeko bistaratu nahi den erregistroaren aukeraketa	57
Irudia 32. RR seinalea edo RR seinale interpolatua bistartzeko aukera	58
Irudia 33. RR seinaleari batezbestekoa kenduta (detrend) lortzen den emaitzaren adibidea ...	58
Irudia 34. PSD grafikoaren adibidea, maiztasun tarte ezberdinak bereiziz	58
Irudia 35. SD2 eta Peak LF parametroen ROC kurbak eta AUC balioak	59
Irudia 36. Lau parametro onen ROC kurba eta haien AUC balioa	60
Irudia 37. Parametro hoberen korrelazio matrizea	61

LABURDUREN ZERRENDA

GKB	Gaixotasun kardiobaskularra
GHI	Garuneko hodietako istripua
FA	Fibrilazio aurikularra
EKG	Elektrokardiograma
FAP	Fibrilazio aurikular paroxistikoa
HRV	Heart Rate Variability
NSP	Nerbio-sistema periferikoa
NSA	Nerbio-sistema autonomoa
NSS	Nerbio-sistema sinpatikoa
NSP	Nerbio-sistema parasinpatikoa
PDE	Potentzia dentsitate espektrala
VLF	Very low frequency
LF	Low frequency
HF	High frequency
GUI	Graphical user interface
TP	True positive
FP	False positive
TN	True negative
FN	False negative
Se	Sentikortasuna
Sp	Berariazkotasuna

1. SARRERA

Gaur egun, gaixotasun kardiobaskularrak (GKB) dira heriotza-kausa nagusia Espainian (heriotza guztien %29,17aren eragile izanik), minbiziaren eta arnas sistemako gaixotasunen gainetik. Euskal Autonomia Erkidegoan (EAE), GKBen heriotza-tasa antzekoa da, urteko %26,98koa izanik [1]. GKBen artean, garuneko hodietako istripuak (GHI) dira heriotza kausa nagusiena.

Ondorioz, osasun publikoarentzat arazo garrantzitsuenetako bat da GHI prebentzioa eta tratamendua, heriotza ugarien eragilea izateaz gain helduen artean ezgaitasun eragile nagusienetako bat baitira eta eskasia neurologikoa eragiten baitute adineko pertsonen artean.

GHI burmuineko edozein ataletara doan odoljarioaren eteteari deritzo. Bi GHI mota bereiz ditzakegu. Alde batetik, odolbatu batek burmuinera doan zain baten odoljarioa eteten duenean GHI iskemikoa dago. Bestetik, zain bat apurtzen denean eta odola burmuinera isurtzeari GHI hemorragikoa deritzo.

GHI iskemikoa jasateko zenbait arrisku faktore azpimarra ditzakegu. Hasteko, bizimoduarekin erlazionatutako arrisku faktoreak daude non gizentasuna eta jarduera fisikoaren falta nagusi diren. Legezkanpoko drogen kontsumoa eta gehiegizko alkohol kontsumoa ere oso kaltegarriak izan daitezke. Horrez gain, badira arrisku faktore aldaezinak; hala nola, familiako aurrekariak izatea, edo adina ere arrisku faktoretzat har daiteke; zenbat eta nagusiagoa izan, orduan eta aukera handiagoa dago GHI sufritzeko. Azkenik, arrisku faktore medikoak ditugu. Hauen barruan arteria-presio altua, diabetesa zein GKBak aurki ditzakegu. Gaixotasun arterial koronarioak, miokardiopatiak, bihotz-gutxietasunak eta fibrilazio aurikularrak (FA) odolbatua eragin dezakete buruko isuria sortuz.

FA arritmia guztien artean bihotzeko erritmo normalaren alteraziorik ohikoena da. Bihotzaren goialdeko bi barrunbeetan (aurikuletan) gertatzen den fibrilazioa da seinale elektriko desantolatuen eraginez. Nahiz eta FA aritmien artean larriena ez izan, berez ez baitu bizitza arriskuan ipintzen, oso kaltegarria da bizitzarako.

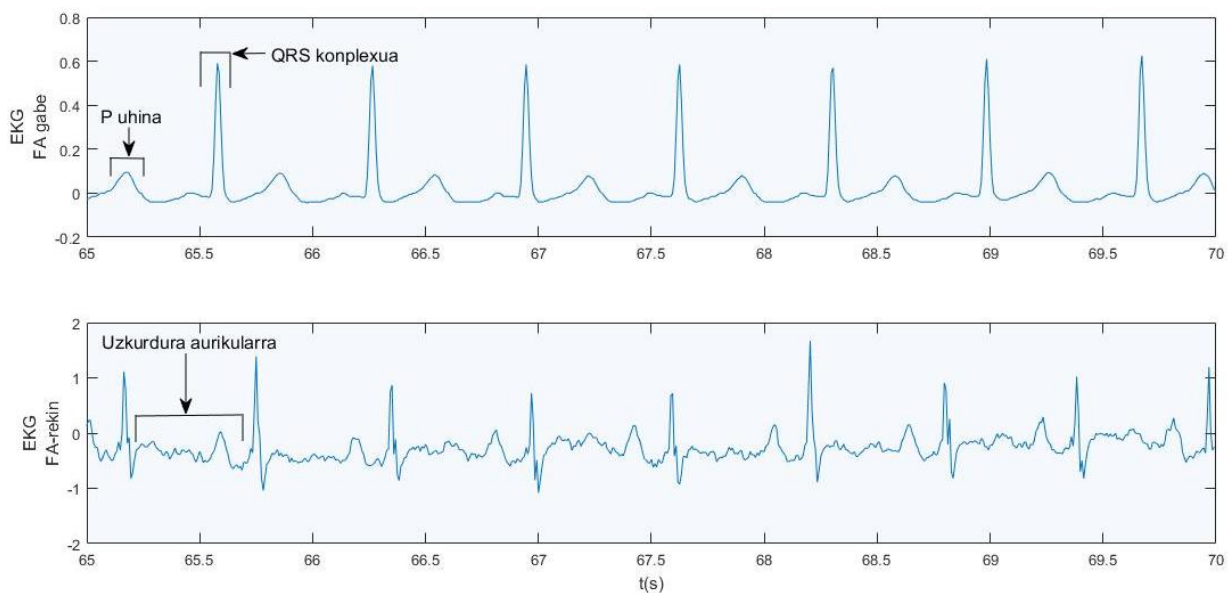
Aurretik aipatu bezala, FA GHI jasatearen arrisku faktore garrantzitsua da. FA daukan pertsona batek biztanleria orokorra baino bost aldiz probabilitate altuagoa dauka GHI izateko [2]. GHI guztien %15en eragilea dela kalkulatu da [3,4], eta GHI iskemikoen %20a [5]. Gainera, gizartearen zahartzearen ondorioz, FAren prebalentzia handituz joango da. FA duten 70 urte goragoko pazienteen GHI-ko prebalentzia bikoiztuz doa hamarkada bakoitzeko [6]. Horrez gain, FArekin erlazionatuta dauden GHI-ak emaitza kaskarragoak sortzen ditu. Normalean, gradu handiagoko morbiditatearekin lotuta dago eta ospitalizazio kostuak garestiagoak izaten dira [7].

Hau guztiagatik, FA auresatea funtsezkoa dela ondoriozta dezakegu, hori bera izango da proiektu honen muina.

2. TESTUINGURUA

Fibrilazio aurikularrak azkeneko hamarkadan garrantzia askoz handiagoa hartu du medikuntza munduan. Azkenengo datuek Espainian FAren prebalentzia 40 urte goragoko biztanlerian %4a baino handiagoa dela erakusten dute. Beraz, gaur egungo, espainiar demografia kontuan hartuta, Espainian milioi bat pertsona baino gehiagok FA dutela pentsa genezake, eta adituek diotenez, gaixotasun honen prebalentzia handituz joango da populazioaren zahartzeak eraginda.

Aurikuletan ematen diren taupadak desordenatuak direnean fibrilazio aurikularra agertu dela esaten da. Bihotzak lau barrunbe ditu; goialdeko bi barrunbeei aurikulak deritze, eskuinekoa eta ezkerrekoa, beheko bi barrunbeak ezkerreko eta eskuineko bentrikuluak direlarik. Orokorrean, bihotzak erregulariki nodo sinusaletik, bihotzaren berezko taupada markagailua-, seinale elektriko bat bidaltzen du eskuineko aurikulara. Gero, seinalea bi aurikuletatik zehar oso azkar hedatzen da aldibereko despolarizazioa eraginez, eta azkenik, uzkurdua bateratu batean bukatzen da. FArekin, seinaleak era guztiz desordenatu batean hedatzen dira aurikuletatik zehar, eta ondorioz, bihotzaren berezko taupada-markagailua baliogabetzen da. Hori dela eta, uzkurdua handia eta bakarra izan beharrean, uzkurdua txiki asko antzematen dira, aurikuletan dardarak eraginez. Elektrokardiograman (EKG) normalean, uzkurdua aurikularra P uhinarekin bat etortzen da, eta honen ostean QRS konplexua dator, azken hau uzkurdua bentrikularra delarik, 1. irudiko goiko grafikan ikusten den bezala.



Irudia 1. Goiko grafikan EKG normala ikus daiteke non P uhina argi eta garbi bereiz dezakegun. Azpiko grafikan, aldiz, FA daukan paziente baten EKG-aren adibidea ikusten da.



Irudia 2. FA-ren eraginez sortzen diren mugimendu desorekatuak

FA daukaten pazienteetan, aurikuletako zonalde txiki asko denboraz kanpo uzurtzen dira, 2. irudian adierazten den bezala. Horren ondorioz, EKGan P uhina izan beharrean, gorabeherak aurkezten dira, non piko bakoitza aurikuletan astintzen ari den puntu batekin bat etortzen den, 1. irudiko azpiko grafikan adierazi den bezala.

Batzuetan, aurikuletan ematen diren seinale txiki hauetako bat bentrikuluetara doa uzurtzura bentrikularra sortaraziz. QRS konplexu hauek, ordea, denbora-tarte irregularretan hedatzen dira, orokorrean 100 eta 175 taupada minutuko bihotz-maiztasun altuarekin. Taupada normal batean, uzurtzura aurikular koordinatu batek “ostikada aurikular” moduan ezaguna den odol kantitate txiki bat eransten du. FA duten pertsonak ostikada aurikular hau galtzen dute, nahiz eta honek ez duen heriotza arriskurik adierazten.

Fibrilazio aurikularra sailkatzeko hainbat proposamen egin dira. Hala ere, American College of Cardiology (ACC), American Heart Association (AHA) eta European Society of Cardiology (ESC) erakundeek gomendatzen dutena, bere sinpletasunean eta garrantzi klinikoa oinarrituz, hurrengoa da:

- **Lehenengo agerpena:** bereizi egin behar da FAren lehenengo agerpena edo beste motako patologia izan den. Dena den, oso zaila da lehen agerpen hau bereiztea; izan ere, ez dago argi zein den bere iraupena.
- **Paroxistikoa:** Aipatu denez, ez da erreza lehenengo agerpena hautematea. Beraz, FA duten pertsonak gertaera paroxistikoeekin hasten direla esan genezake. Iraupen motzeko agerpenak dira; segundo batzuetatik minutu edota ordu batzuetako iraupenezkoak, eta ez dira jarraian errepikatzen 7 eguneko baino tarte luzeago batean. Izan ere, ehuna oraindik nahiko osasuntsu dago. Hala ere, denboraldi luzeagoko gertaera paroxistikoko errepikatuek aurikuletako zelulak gero eta gehiago estresatzea eragin dezakete. Denbora igaro ahala, estres hori dela eta, aurikuletako zelulek fibrosi progresiboa jasan dezakete. Honakoa ematen denean, FA gertaera ezin da modu espontaneoan moztu, eta beraz, pazienteek FA etengabea daukate.
- **Etengabea:** aste bateko baino gehiagoko iraupena izaten du. Kasu honetan, arritmia medikazio bitartez edota kardiobertsio elektriko bitartez mozten da.
- **Iraunkorra:** medikuaren eta pazientearen arteko erabakiagatik arritmiaren lehengoratzea saiateren ez denean edo saiaturako terapeutikarekin arritmia hobetzen ez denean FA iraunkorra kontsideratzen da.

Proiektu hau, 2001.urtean planteatu zen “Predicting the Onset of Paroxysmal Atrial Fibrillation” izeneko erronkan oinarritzen da. Hortaz, fibrilazio aurikular paroxistikoa (FAP) gauzatu baino segundo batzuk lehenago EKG-n agertzen diren aldaketak desberdintzen saiaturako gara gure helburua arritmia aurreratuko algoritmo fidagarria garatzea izango delarik. Izan ere, gaur egun, ez da FAP detektatzeko fidagarria den metodorik baliotu.

Erronkarekin hasteko, PhysioNet bidez eskuragarri dagoen datu-basea erabili da. Datu-baseak pertsona osasuntsu eta FAP duten pazienteen kasuak biltzen ditu. Epe luzeko EKG erregistroetan oinarritzen da, non erregistro bakoitzeko informazioa bi kanal ezberdinetatik lortu den. Erregistroak kopuru bereko bi multzotan bereizten dira. Batetik, ikasketa multzoa dago, eta bestetik, test multzoa. Datu-baseak, digitalizatutako EKG seinaleak (guztien laginketa maiztasuna 128Hz-koa izanik) eta taupaden anotazioen informazioa barne hartzen ditu.

Ikasketa multzo barruan, EKG seinale bakoitzak 30 minutu irauten du. QRS anotazioen informazioa erabiliz, RR tarteak [8] kalkulatu dira. Ondoren, denbora-domeinuko zein frekuentzia-domeinuko parametro ezberdinen balioak lortuko dira. Aipatutako informazio guzti hori konbinatuz, FAP aurrerako metodoak garatu dira, eta azkenik, metodo horiek estatistikoki ezaugarritu eta test datu-basean azken emaitzak lortuko dira.

3. HELBURUAK

3.1. HELBURU NAGUSIA

Proiektu honen helburu nagusia FAP auresateko metodoa lortzea da. Horretarako, ezinbestekoa izango dugu metodoa behar bezala garatzea hurrengo bi kasu ezberdinak bereizi ahal izateko.

Batetik, metodoak FAP duten pazienteak eta paziente normalak bereizteko gai izan behar du. Aipatuenez, gure datu-basean paziente osasuntsu eta FAP duten pazienteen erregistroak izango ditugu. Beraz, hauen EKG seinaleen ezberdintasunetan oinarrituz bi paziente mota hauek bereizteko algoritmoa garatzen saiatuko gara.

Bestetik, FAP duten pazienteen artean, honako fibrilazioa noiz duten edo ez bereiztea izango da emango dugun hurrengo urratsa. Horretarako, FAP duten pazienteen artean, pazienteak gertaera paroxistiko bat jasan aurretiko EKG seinaleak eta gertaera normal batean lortutako EKG seinaleak konbinatuz lan egingo da.

Azken helburu hau lortzeko, aurretik beste tarteko helburu partzial batzuk bete beharko dira.

3.2. HELBURU PARTZIALAK

3.2.1. DATU-BASEA DEFINITU

Proiektu honetan finkatutako helburu teknikoa lortzeko eman behar den lehenengo pausua datu-base bat lortzea da. Datu-base horrek bereizi nahi diren pazienteen (FA duten eta paziente osasuntsuen) EKGak izango ditu, eta EKG hauek mediku espezialistek anotatuta egon behar dute. Bestetik, komeni da baita ere gertatutako taupaden (QRS konplexuen) markak izatea, hauek baitira FA eta FAP duten pazienteen bereizgarrietako bat.

3.2.2. SEINALEAK BISTARATZEKO TRESNAK GARATU

Behin aipatutako datu-basea lortuta, oso garrantzitsua da gordetako seinaleak bistaratu ahal izateko tresnak garatzea. Izan ere, proiektu honetan EKG seinale askorekin lan egingo da. Beraz, seinale horietatik ateratako parametroen balioak, eta horren ondoren, erabiliko diren algoritmoak garatzeko tresna dinamikoen erabilpena ezinbestekoa da.

3.2.3. ALGORITMOA GARATU

Esan bezala, proiektu honen helburu nagusia FAP auresateko algoritmo bat garatzea eta ezaugarritzea da. Horretarako, ezinbestekoa izango da EKtik abiatuta auresate parametroak diseinatzea eta probatzea, datu-basean anotatutako bi taldeak bereizteko balio dutenak. Parametroak diseinatu ondoren, parametro horiek bi taldeetan dituzten desberdintasun estatistikoak baliatuko dituzten ikasketa automatikoko algoritmoak erabiliko dira auresate gaitasuna handitzeko.

3.2.4. EMAITZEN AZTERKETA ESTATISTIKOA BURUTU

Algoritmo horren zehaztasun diagnostikoa erabakitzeke metodologia eta balio esanguratsuak definituko dira eta algoritmoaren erabakitze gaitasuna modu estatistikoan ezaugarrituko da. Horretarako, algoritmoaren bitartez lortutako emaitzak fitxategi batean gordeko dira. Horiekin batera, erreferentziako balioak gordeko dira datu-base berean. Horrela, aplikatutako metodoarekin sortzen den errorea kalkulatu da.

4. ONURAK

Proiektu honekin hainbat onura lortuko dira eremu ezberdinetan, garrantzitsuenak eremu teknikoa, ekonomikoa eta soziala direlarik.

4.1. ONURA SOZIALA

Gaur egun, FA gehien ematen den arritmia da. Ez da arriskutsuena, berez ez duelako heriotzera eramaten, baina oso kaltegarria dela esan daiteke. Aurretik aipatu den bezala, FA daukan pertsona batek GHI sufritzeko probabilitate altuagoa dauka. Gainera, urteak aurrera joan ahala, gehiago dira arritmia hau sufritzen duten pertsonak. Horren ondorioz, oso garrantzitsua da FA garaiz detektatzeko metodoak edukitzea.

Proiektu honetan FA aurreikusteko metodo bat aurkezten da, metodo honekin posible izango litzateke FA gertatuko dela antzematea hau gertatu aurretik. Garaiz detektatzea lortu ezker, posible izango da jada existitzen diren tratamenduak erabiliz arritmia mota horri erantzuna ematea. Tratamendu hauen helburua bihotzaren erritmo desorekatuen agerpena saihestea da, eta hau ematen den egoeratan, bihotz-maiztasuna kontrolatzen saiatzen dira zailtasunak urritzeko asmoz (enboliak, adibidez).

4.2. ONURA TEKNIKOAK

Garatutako baliabide eta tresnak izan dira proiektu honen onura tekniko esanguratsuenak. Batetik, EKG datu-basetik irakurtzeko tresnak garatuko dira, eta tresna hauek beste proiektu batzuetarako erabilgarriak izan daitezke. Bestetik, sortutako datu-basean, seinale horiez gain, beste seinale batzuen informazioa batzen da; hala nola, EKG seinaleen informazioa. Hortaz, datu-base hau seinale hauen ingurunean etorkizunean eginiko hainbat ikerketetan lagungarria izan liteke.

Horrez gain, proiektu honetan proposatuko den algoritmoa medikuntza arloan FA auresateko metodo bezala erabilgarria izan daiteke.

4.3. ONURA EKONOMIKOAK

Aurretik aipatu den bezala, FA da osasun sisteman kosturik handien eragiten duen arritmia. Beraz, proiektu honekin arritmia honekin erlazionatuta dauden kostuak gutxitzea posiblea izango litzateke. Aurkeztutako algoritmoarekin modu merke batean beste metodo batzuekin lortzen diren emaitza antzekoak eskuratzea posible izango da.

Honekin guztiarekin zera lortuko da, FA kasu baten aurrean emaitza hobea eta azkarragoa ematea, horrela arritmia mota honek eragiten dituen alboeraginak arinduz, eta beraz, osasun sistemak paziente horietan gastatu beharreko diru kopurua gutxituz.

5. ARTEAREN EGOERA

FA bihotz-arritmiarik ohikoena da. Gehienetan, arteria-presio altua duten edota bihotzean beste gaixotasun motaren bat duten pertsonen agertzen zaie (balbuletako gaixotasunak dutenei edo bihotzekoa jasaten dutenei, adibidez). Horren ondorioz, gaixotasun honek bihotz-gutxiegitasunagatik edo enbolia batengatik morbiditate arrisku altua zein bat-bateko heriotza (BBH) eragin ditzake.

Hainbat ikasketek Faren prebalentziaren balioespenak egin dituzte munduko hainbat herrialdeetan. Framingham ikasketaren arabera, 50-59 urte bitarteko pertsonen artean Faren prebalentzia %0.5ekoa da, eta 80-89 urte bitarteko pertsonena, ordea, %8.8 [9]. Bestalde, Sudlow eta bere ikerkuntzako lankideek onetsi zuten FAK prebalentzia altuagoa izan dezakeela Erresuma Batuan Estatu Batuetan baino. Horrez gain, Britainia Handian 65 urteko goragoko pazienteen artean gutxienez 160.000-k FA izan ahal luketela aditzera eman zuten [10]. Hala ere, eginko balioespen gehienak FA iraunkorrari aplikatuta daude, FAPri erlazionatutako datuak urriagoak direlarik.

Medikuek lehen mailako atentzioan zein ospitaletan antzemandako Faren kasu osoen %25 eta %62 bitartekoak FAP direla onesten dute [11-13]. (Prebalentziari buruz bildutako informazioa oso aldakorra da herrialde ezberdinen artean).

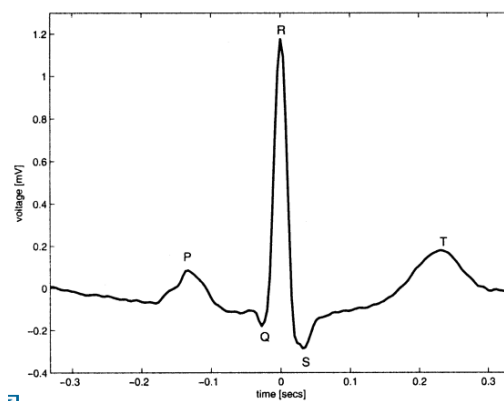
FAP-ko gertaerak; hots, modu espontaneoan amaitzen direnak, FA etengabearen gertaeren aurrekariak dira. FAP kasuen %18a FA etengabera bilakatzen da lau urte igaro ondoren [14]. FA etengabearen gertaerak, ebaketa farmakologiko, elektriko edo kirurgikoko aplikazioen bitartez soilik mozten dira. Horren ondorioz, FAPren auresateak gaixotasun honekin lotuta dauden kostu klinikoak gutxituko lituzke eta pazienteen bizi-kalitatea hobetuko litzateke.

FA pazienteen tratamendua oraindik ez da guztiz ona. Izan ere, arritmia mota honek izaera mailakatua izateagatik Faren gertaeren hasiera eta amaierak azalpenik gabe ematen dira. Hainbat autorek metodo ezberdinak garatu dituzte FA auresateko EKG erregistroen informaziotik abiatuta.

5.1. EKG SEINALEA

EKG denboran aldakorra den seinalea da, bihotzeko zuntz-muskuluak uzurtzea eta gero erlaxatzea eragiten duen korrontearen fluxua islatzen duelarik. EKG azalaren gainazalean kokatutako bi elektrodoen arteko potentzial diferentzia erregistratzean lortzen da. EKGren ziklo normal bakoitzak bihotzeko taupada bakoitzean gertatzen den ondoz ondoko despolarizazio/birpolarizazio aurikularra eta geroago despolarizazio/birpolarizazio bentrrikular prozesuak azaltzen ditu. Azken finean, EKG bihotzaren egoera era ez inbasiboan aztertzeke erabiltzen den froga fisiko bat da.

EKG batek ondorengo ezaugarriak ditu bihotz taupada arrunt bat erregistratzen denean: P uhina, QRS konplexua eta T uhina, 3. irudian ikus daitekeen moduan.



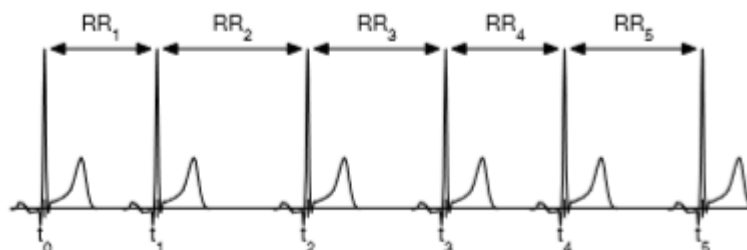
Irudia 3. EKG seinalea

Beraz, EKG ondorengo ataletan zatituta dago:

- **P uhina:** despolarizazio aurikularrari dagokion seinale elektrikoa da. P uhinaren hasierari eskuin aurikularen despolarizazioa deritzo, P uhinaren bukaera, ordea, ezker aurikularen gainjartzearen emaitza delarik. P uhinaren birpolarizazioa, despolarizazio bentrikularragatik (QRS konplexua) estalita geratzen da. P uhinaren amaiera eta QRS konplexuaren arteko denbora-tarteari PQ-tartea deritzo.
- **QRS konplexua:** esker eta eskuineko bentrikuluen uzkuetzea eragiten duen korrante elektrikoari, -despolarizazio bentrikularrari-, dagokio. Seinale hau aurikulek sortutako seinalea baino askoz indartsuagoa da. Izan ere, atal honetan parte hartzen duten gihar masak handiagoak dira. Beraz, EKG seinaleko anplitude handiagoko tartea da.
- **T uhina:** bentrikuluen birpolarizazioa adierazten duen EKG seinalearen zatia da.

Azkeneko bi atal hauekin erlazionatuta dauden bi tartek existitzen dira. Batetik, QT-tartea. Bentrikuluen despolarizazio eta birpolarizazioari dagokio, eta QRS konplexuaren hasieratik T uhinaren amaiera arte neurtzen da. Bestetik, ST-tartea dago.

Aipatutako EKG seinalearen atal hauek kontuan hartuz, Heart Rate Variability (HRV) deituriko analisia aplikatu daiteke. HRV analisia bihotz erritmoa denbora tartek luzeekin (5 min – 48 ordu arte) erlazionatuta dago. Izan ere, analisi honen bidez taupaden arteko tarteen aldakortasuna neurtzen baita, gorputzaren hainbat mekanismo fisiologikoekin lotuta dagoena. Horretarako, bihotz erritmoa uneoro neurtzen da taupaden iraupena neurtuz, eta horretarako, EKGko punturik esanguratsuena hartzen da, R uhina alegia. Horrela, bihotz taupada baten iraupena elkarrekiko jarraiak diren bi R uhinen arteko tartea da, RR tartea. Hau hobeto uler dadin 4. irudia gehitu da.



Irudia 4. EKG seinaletik lortutako RR tartekak

Behin tarteak kalkulatu HRV analisia egiten da, analisi honek hainbat parametroen kalkulua egitea ahalbidetzen du, arritmiekin erlazionatutako arazoak antzemateko lagungarriak izan daitezkeenak.

5.2. HEART RATE VARIABILITY

Taupada bakoitzarekin lortutako RR tarteen bitartez bihotz maiztasuneko denbora seriean egindako aldaketan analisiari HRV analisia deritzo. HRV analisiak GKBen ebaluazioa eskaintzen duela frogatu da [15].

Izakien nerbio-sistema bi zati nagusitan banatzen da. Batetik, nerbio-sistema periferikoa (NSP) dago. Honek kanpoko informazioa jasotzen du eta nerbio-sistema zentralera eramaten du. Horrez gain, hark emandako aginduak muskuluei igortzen die, azken hauek behar den erantzuna eman ahal izateko. Bestetik, nerbio-sistema autonomoa (NSA) dago. Bigarren hau, gorputzaren barne-organoen funtzioa erregulatzear arduratzen da, funtzio automatikoak, -nahi gabe egindakoak-, betetzen dituen nerbio-multzoa delarik.

NSA bitan banatzen da: nerbio-sistema sinpatikoa (NSS) eta nerbio-sistema parasinpatikoa (NSP). NSS energia gastua eskatzen duten ekintzekin erlazionatuta dago. NSP, ordea, energia gordetzearen arduraduna da. Bata bestearen antagonista da. NSS eta NSPen arteko efektuen oreka taupada bakoitzeko bihotz-zikloen aldaketetan islatzen dela pentsa ohi da [16].

Behin QRS konplexua zein denbora unetan gertatu diren antzemanda, HRV denbora seriea lor daiteke. Izan ere, taupaden arteko interbalok; hau da, RR tarteak, R uhinaren ondoz ondoko denboren arteko diferentzia kalkulatu lortzen dira. Beraz, RR tarteekin lortutako denbora seriean laginen arteko distantzia ez da konstantea, eta hortaz, denboraren menpe adierazten da.

RR tarteak behin lortuta, metodoen analisia egitera pasatzen da. Domeinu ezberdineko parametroen kalkulua egingo da, betiere kalkulu hauek bost minutuko EKG seinaleetan oinarrituta, denbora laburreko HRV analisia alegia.

- **Denbora-domeinuko metodoak:**

Denbora-eremuko metodoak burutzeko errezenak diren metodoak dira. Izan ere, metodo hau RR tarteko ondoz ondoko balioei zuzenean aplikatzen zaie. Hauen barruan zenbait parametro desberdinu ditzakegu:

- **SDNN:** Parametro honek RR tarteen desbideratze estandarra adierazten du.
- **RMSSD:** Ondoz ondoko RR tarteen arteko desberdintasun koadratikoaren erro karratua aplikatuz lortutako parametroa da.
- **pNN50:** RR tartearen seriearen barruan, 50 ms baino gehiagoko desberdintasuna daukaten ondoz ondoko lagin kopuruari NN50 deituriko parametroan islatzen da. Parametro hau jarduera parasinpatikoa ebaluatzeko proposatu da. Zenbaki hori serie osoan dauden lagin kopuruarekin zatitzean pNN50ren balioa lortzen da.

- **Maiztasun-domeinuko metodoak:**

Maiztasun-domeinuko metodoak aplikatu ahal izateko RR tartearen potentzia dentsitate espektrala (PDE) kalkulatu behar da. Horretarako, beharrezkoa da laginen arteko distantzia konstantea izatea. Beraz, denbora-domeinuan ez bezala, kasu honetan, interpolazioa deituriko prozesua egin behar da, Fourieren transformatua (FFT) aplikatu aurretik.

HRV erregistroetan erabiltzen diren ohiko frekuentzia bandak hurrengoak dira: batetik, oso maiztasun baxuko banda (VLF, ingeleseko “very low frequency”), 0-0.04 Hz-eko tartea hartzen duena. Bestetik, maiztasun baxuko tartea (LF, ingeleseko “low frequency”), 0.04-0.15 Hz bitartekoa. Azkenik, maiztasun altuko banda (HF, “high frequency”), 0.15-0.4 Hz-ko maiztasun tartea.

Aipatutako maiztasun tarte hauekin erlazionatuta dauden eta analisirako erabiliko diren parametroak hurrengoak dira:

- **Peak frequencies:** VLF, LF eta HF maiztasun tarte bakoitzeko potentzia altuena duen maiztasun balioak dira, Hertziotan (Hz) neurtuta.
- **Absolute power:** VLF, LF eta HF banda frekuentzien potentzia absolutua da. Hau da, frekuentzia banda bakoitzaren barruan dagoen RR tarte osoaren potentzia espektralaren kalkulua egiten da. Balio hauek, ms^2 -tan neurtzen dira.
- **Relative power:** aurreko atalean maiztasun banda bakoitzeko lortutako potentzia absolutuak potentzia totalarekin zatitzen badira, banda bakoitzeko potentzia erlatiboa lortzen da, ehunekotan adierazten delarik.
- **Normalized power:** LF eta HF banden potentzia unitate normalizatueta adierazten du.
- **LF/HF :** LF eta HF banden arteko potentzia erlazioa.

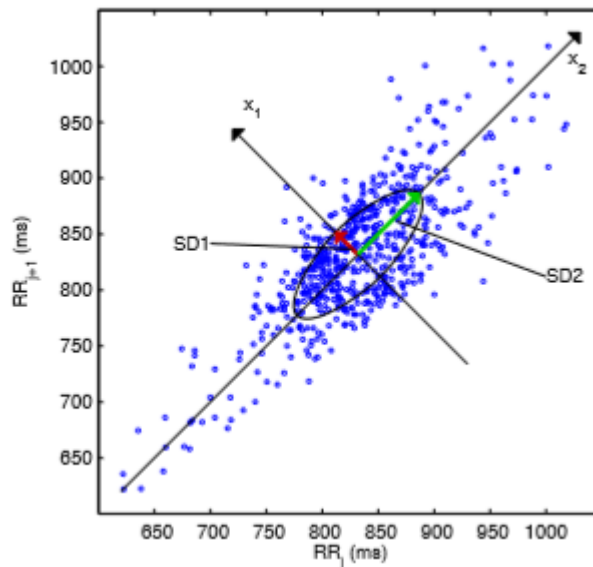
- **Metodo ez-linealak**

Analisi honek emaitza zuzenak eman ahal izateko, ez da nahikoa denbora- eta maiztasun-domeinuko parametroak soilik kontuan hartzea. Beraz, ez-linealak diren metodoak erabili beharko dira. Honen barruan, zenbait parametro aurki daitezke:

- **Poincaré plot:**
Ondoz ondoko RR tarteen arteko korrelazioaren irudikapen grafikoa da. Grafika honen itxura da metodo honen funtsezko ezaugarria. Grafikoa irudikatzen duten puntuak elipsea osatzen dute (hurrengo ataletan ikusiko da).

Metodo honekin, irudikapen grafikoaz gain, bi parametro esanguratsu lortzen dira. Batetik, SD1 parametroa. Hau, lerro identitatearekiko perpendikularrak diren puntuen desbideratze estandarri dagokio. Bestetik, SD2 parametroak, lerro identitatearekiko paraleloak diren puntuen desbideratze estandarri adierazten du. Aipatutako bi parametroak denbora-domeinuko zenbait parametroekin erlazionatuta daude.

5. irudian, azaldutako grafikoaren adibidea ikus daiteke. Aipatu denez, elipse itxurakoa da, eta bertan SD1 eta SD2 parametroak bereiz daitezke.



Irudia 5. RR tarte jakin batekin lortutako Poincaré Plot

- **Approximate entropy:**
Entropia hurbildua (ApEn) izeneko parametroak seinalearen konplexutasuna edo erregularitatea neurtzen du. ApEn-en balio altuak seinalea oso irregularra dela adierazten du. Era berean, ApEn balio baxuak izateak seinalea normalagoa dela esan nahi du.
- **Sample entropy:**
Laginen entropia (SampEn) deituriko parametro honek antzeko garaien denbora serieak aztertzen ditu eta balio ez-negatibo bat esleitzen dio konplexutasun handiagoa edo datuetan erregularitateak dituzten balio handiagoko sekuentziei.
- **Fuzzy entropy:**
FuzzyEn deituriko parametroak entropia nahastua adierazten du.

5.3. SAILKATZAILEAK

FA auresateko metodoa lortzeko prozesua burutzeko, egin beharreko lehenengoko pausua aurreko atalean azaldutako parametroen balioak lortzea da. Behin balio horien kalkulua egin ondoren, parametro horiek konbinatuz, algoritmoa lortuko litzateke. Azkeneko pausu honi, proiektu honetan, sailkatzailearen ideia aplikatzeari dagokio.

Horretarako, urrats batzuk jarraitu behar dira. Hasteko, algoritmoa garatu ahal izateko erabiliko diren parametroak finkatu behar dira. Parametro horiek, aurreko atalean azaldutakoak direlarik, FAP duten pazienteen eta paziente osasuntsuen datuetatik lortuak dira. Beraz,

parametro horiez gain, hori bera islatzen duen beste aldagai bat (FAP duten ala ez) erabili beharko da algoritmoa garatzeko. Hori guztia erlazionatuz, eta algoritmo ezberdinak aplikatuz sailkatzaile ezberdinak lortuko dira. Azkenik, sailkatzaile hori datu berriei aplikatuz, paziente berri hauek arritmia duten ala ez auresatea da helburua. Hala ere, emaitzak aztertuz, FA auresateko metodoa zein neurritan fidagarria den adierazi beharko da, eta erabilitako parametroak zuzenak diren ala ez jakitera emango duen analisisa egin beharko da.

Azken finean, hau da medikuntzan jarraitu ahal daitekeen prozesua arritmia bat antzeman ahal izateko. Hala ere, beste zenbait metodo erabil daitezke FA auresateko.

5.4. FA AURESATEKO METODO EZBERDINAK

Aurretik esan bezala, EKG seinalearen bitartez, eta honekin HRV analisisa aplikatuz, FA auresateko metodoa garatu daiteke. EKG seinalea lortzen duten bi gailu erabili ohi dira medikuntza arloan. Batetik, elektrokardiografoa dago, gaur egun kardiologian eta medikuntza orokorreko diagnostikoan oinarritzko tresna delarik. Bestalde, Holter izeneko proba existitzen da, Holter gailuaren bitartez, 6. irudian ikusten den moduan. Kasu honetan ere, pazienteak gainazalean elektrodoak izango ditu, baina hauen informazioa honek daraman monitore batean erregistratuko da. Normalean, proba hau 24 orduz egiten da. Beraz, pazienteari buruz lortutako erregistroak luzeagoak dira informazio gehiago emanez.



Irudia 6. Holter proba gauzatzeko gailua

Hala ere, arritmia ezberdinak antzemateko ekokardiograma deituriko prozesua egin daiteke. Metodo honek ultrasoinuak erabiltzen ditu bihotzaren egitura bistaratzeko, eta horren bitartez, bihotzak odola bonbeatzeko duen gaitasuna aztertu ahal izateko. FA duen paziente batean, adibidez, aurikuletan dardarak gertatzen direla ikusiko litzateke eta ez hauen uzkurdura.

6. METODOLOGIA

Proiektu honen xede nagusia FA aurreratzeko metodoa garatzea izan da. Horretarako, zenbait pausu jarraitu behar izan dira. Beraz, atal honetan, aurrera emandako pausu horiek zehaztasunez adieraziko dira.

6.1. DATU-BASEA SORTZEAREN PROZESUA

Lehenengo fase honetan bi prozesu bereiziko dira. Hasteko, Physionet-etik oinarritzat hartu den datu-basea azalduko da, eta honen ondoren, proiektuan zehar erabili den datu-basea zehaztasunez erakutsiko da.

6.1.1. OINARRIZKO DATU-BASEAREN AZALPENEA

Proiektu hau Physionet-ek proposatutako erronka bat betetzean datza [17], erronka hau FAP aurreratzeko metodoa lortzeko gaitasuna erakustean oinarritzen delarik.

Hasteko, Physionet-ek webgunean erronka betetzeko eskuragarri dituen datuak [18] ordenagailura jaitsi dira. Eskuratutako datu-base honetan, Holter probaren bidez lortutako EKG seinaleen erregistroak gordeta daude, 128Hz-ko maiztasunarekin laginduta daudelarik. Horrez gain, QRS anotazioen denbora unen informazioa eskaintzen da.

Erregistroak bi multzo nagusitan bananduta daude: batetik, garapenezko multzoa dago, eta bestetik, test-erako multzoa. Garapenezko multzoaren barruan, erregistro bakoitza paziente osasuntsu bati edota FAP duen paziente bati dagokion adierazten da. Alabaina, test-erako multzoan, erregistro bakoitza ze pazienterena den adierazten duen informazioa ezezaguna da. Hortaz, FA aurreratzeko metodoa garatu ahal izateko, garapenezko multzo barruan dauden erregistroekin lan egingo da.

Garapenezko multzoan guztira ehun pazienteren informazioa aurki daiteke, paziente horietariko 50 osasuntsuak direlarik, eta beste 50ak PAF duten pazienteak. Paziente bakoitzak lau erregistro ditu, eta hurrengo erara egituratuta daude:

- Alde batetik, paziente osasuntsuen EKG normala izango dugu, eta paziente bakoitzeko 30 minutuko bi erregistro normal izango ditugu. Gainera, erregistro horietako bakoitzari, 5 minutuko erregistro bat jarraitzen dio. Kasu honetako bi erregistro multzoek denbora une ezberdinetan ateratako EKG seinaleak biltzen dituzte, baina bata bestearen artean ez da desberdintasun handirik igarriko. PAF pazienteen artean, ordea, bereiztasunak aurkituko dira, orain azalduko den moduan.
- Bestetik, FAP duten pazienteen erregistroak ere binaka multzokatuko dira, aurreko kasuan gertatu den moduan. Lehenengoa 30 minutukoa, eta honen jarraipena 5 minutukoa. Baina orain, lehenengo eta bigarren multzoko erregistroak desberdinak izango dira:
 - Lehenengo multzoa: Hasteko, 30 minututako erregistroan paziente horrek FAP gertaera bat jasateko urrun dagoen une jakin batean lortutako seinalea gorde da. Hortaz, honen jarraipen erregistroak pazientea 5 minutu horietan FAP gertaerarik ez duela ziurtatzen du.

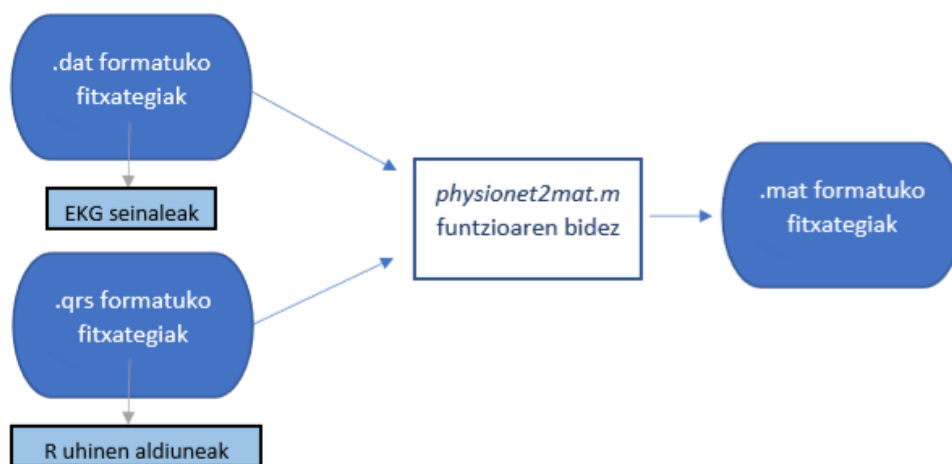
- Bigarren multzoa: Kasu honetan, 30 minututako erregistroan pazienteak FAP gertaera jasotzeke dagoen unean lortutako seinalea gorde da, eta ondorioz, honen jarraipen erregistroan pazienteak FAP gertaera jasaten ari duela islatzen da.

Hau guztia azaldu eta gero, eta proiektuak FA aurrerako metodo bat garatzearen helburua duela kontuan izanik, 30 minutuko seinaleekin soilik lan egingo da. Izan ere, garrantzia emango baitaio FAP gertatu aurreko EKG seinaleen bereiztasuna aurkitzeari, eta ez hau bera gertatzen ari den uneari.

6.1.2. PROIEKTUAN ERABILIKO DEN DATU-BASEAREN GARAPENA

Behin proiektuaren helburua gauzatzeko beharrezkoak diren datuak izanda, datu horiek guztiak datu-base batean gordeko dira. Horretarako, lehenik Physionet formatuan dauden datuak Matlab formatuan gorde beharko dira.

Physionet-etik lortutako datu-base honetan, alde batetik, *.dat* formatuan dauden fitxategiak aurki daitezke, fitxategi hauek seinaleak dituzten artxibategiak direlarik, EKG seinaleak, alegia. Bestetik, *.qrs* formatuko fitxategiak ere badaude, zeintzuk gertaeren informazioa dakartzate; hau da, R uhinen aldiuneak. Hortaz, bi formatu hauetan bereiz daitezkeen fitxategien bihurketa egin da *physionet2mat.m* funtzioa garatuz, hurrengo irudian ikus daitezkeen moduan.

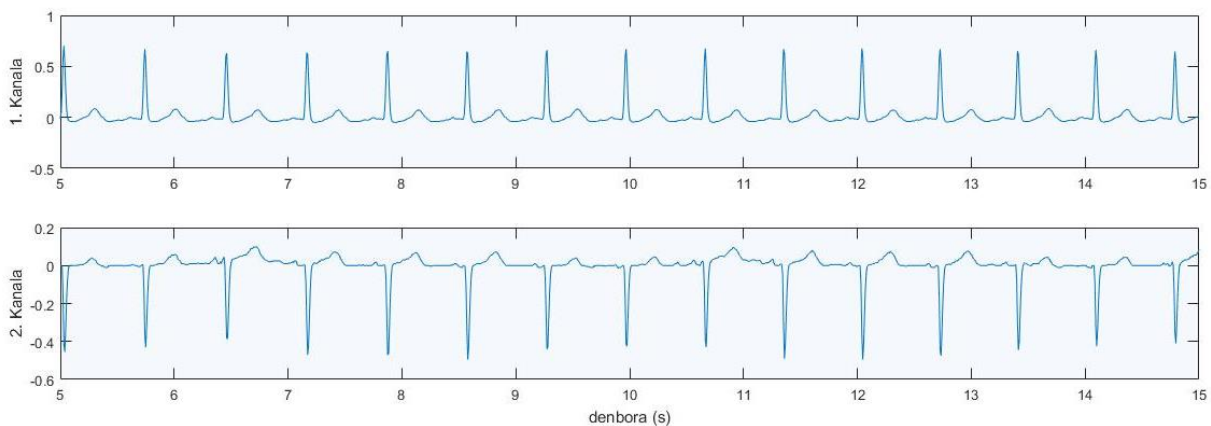


Irudia 7. Jatorrizko fitxategiek jarraitutako bihurketa prozedura

Datuak Matlabeko formatura bihurtu ondoren, hauek banandu ditugu. Gertaerak dituzten fitxategiak, -R uhinen aldiuneak-, izango dira geroago FA aurrerako metodoa lortzeko behar ditugunak. Izan ere, gertaera hauek dira RR tarteak lortzeko erabili direnak. Hortaz, hauen informazioa *metadata* deituriko egitura batean gordeko da, erregistro bakoitzari buruz ezinbestekoa den informazioa gehituz. Honako hau izango da *metadata* barruan gordeko den informazioa:

- Erregistro bakoitzak izen ezberdin bat izango du, elkarren artean ondo bereizi ahal izateko.
- Paziente ezberdinak bereiztu ahal izateko, bakoitzak identifikatzaile bat izango du. Hau da, paziente bakoitza zenbaki bati esleituta egongo da.
- Erregistroa, garapen edo test multzoan dagoen adierazi beharko da. Geroago, garapen multzoarekin soilik lan egingo da, beraz, test barruko multzoan dauden erregistro guztiak baztertuko dira.
- Erregistro bakoitza beste baten jarraipena den ala ez adierazten duen eremua ere izango dugu. Kasu honetan ere, beste baten jarraipenak diren erregistroak ez dira erabiliko.
- QRS anotazioen denbora uneak gordeko dira, hauek EKG seinalearen R uhinen uneekin bat etortzen direlarik.
- Erregistro bakoitzaren taupada mota adieraziko da. Horrela, paziente osasuntsuak FAP duten pazienteetatik bereiztu ahalko da.

Esan dugunez, seinaleak dituzten datuak gainerako datuetatik bereiziko dira. Izan ere, horrela lortuko dena datuen bistaritze eta kalkuluak egiteko prozedurak era dinamikoago batean egitea da. Beraz, proiekturako erabiliko den datu-basean, *metadata* egituraz gain, erregistro bakoitzak bere EKG seinaleen informazioa duten fitxategia izango du. EKG seinale bakoitza bi kanaletan bereizten da, bakoitzak gorputzaren gune ezberdin batetik ateratako informazioa duelarik. Seinale horiek 30 minutukoak dira eta 128 Hz-ko maiztasunarekin laginduta izan dira, beraz, guztiek luzera berdinekoak izango dira, 230.400 laginekoak alegia. Aldagai hau oso erabilgarria izango da erregistro bakoitzaren EKG seinalearen itxura bistaratu ahal izateko. 8. irudian seinale horren adibidea ikus daiteke denbora tarte jakin batean.



Irudia 8. Erregistro jakin baten EKG seinalearen bi kanalen irudikapena 10 segundoko denbora tartean

Erregistro bakoitzaren informazioa era atseginago batean bistaratu ahal izateko, zenbait erabiltzailearen interfaze grafikoaren (GUI, ingeleseko Graphical User Interface) erabilpena ezinbesteko da. Proiektu honetan, bi GUI sortu dira. Lehenengoa datu-basetik lortutako informazioa bistaratzeko balio duena, eta bigarrena, informazio horrekin sortutako RR tarteak bistaratzeko eta tarteetatik abiatuta kalkuluak eta emaitzak bistaratzeko balio duena (I. Eranskina).

6.2. BISTARATZEKO TRESNAK

Bigarren fase honetan, beraz, bistaratzeko GUI tresnen sorketa egin da. Bi izan dira garatu direnak, eta azpiatal honetan bakoitzak erakusten duen informazioa eta honen erabilera azalduko da.

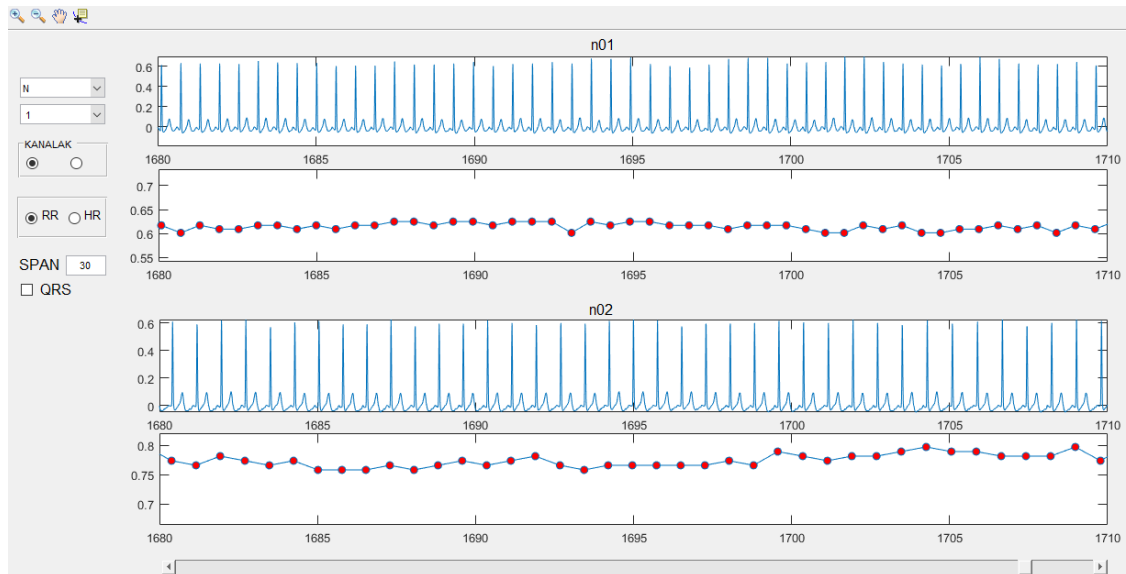
Hasteko, paziente bakoitzeko ditugun EKG seinaleen irudikapena egiteko bistaratzeko GUI den tresna garatu da. Honek, era guztiz dinamiko batean paziente osasuntsuen eta FAP duten pazienteen EKG seinaleen artean aurki daitezkeen zenbait ezaugarri bereiztea ahalbidetzen digu.

Horrez gain, datu-basean gordetako QRS konplexuaren aldiuneak, -R uhinaren uneak, zehatzago adierazteko-, RR tartearen kalkulua egin da. Beraz, EKG seinale bakoitzaren azpian, seinale horri dagokion RR tarteen bistaraketa egin da, horrela, ondoz ondoko taupaden artean dagoen aldea antzeman daiteke. Era berean, RR tarteen balio hauek erabiliz, minutu bakoitzean izandako taupada kopuruaren kalkulua egingo da (HR, ingeleseko "Heart Rate"), eta honako hau bistaratzeko aukera izango dugu.

Hortaz, aipatutako seinale ezberdinak bistaratzeko sortutako GUIk ondorengo ezaugarriak izango ditu:

- Bi menu irristakor izango ditu pazienteak bereizteko. Lehenengoko menua, paziente osasuntsu baten edota FAP duen paziente baten informazioa erakusteko erabiliko da. Horren ondoren, bigarren menuarekin, talde horren barruan eta paziente horrek duen identifikazio zenbakiaren bidez paziente jakin horri buruz daukagun informazioa bistaratuko da.
- Tresnak modu dinamikoan ahalbidetuko du bistaratze tartearen aldaketa (SPAN), QRS anotazioen bistaraketa, kanalaren aukeraketa edota bistaratutako denbora tartean mugitzea (slider herraminta).
- EKG seinale bakoitzaren azpian, RR edota HR tartea bistaratzeko bi botoi bereiziko dira. Horrela, kasu bakoitzean gehien interesatzen zaigun seinalea bistaratzeko aukera izango dugu.

GUIren itxura 9. irudian erakusten da, non aipatutako ezaugarriak antzematen diren. Aurretik esandakoa laburbilduz, goiko ezkerreko aldean, pazientea aukeratuko da. Horren ondoren, paziente horri dagozkien bi erregistroak bistaratuko dira, EKG seinale bakoitzaren kanal bakarra irudikatzen delarik. Horrez gain, RR tarteen kalkulua egingo da, eta hau bera irudikatu dagokion EKG seinalearen azpian. Azkenik, hainbat funtzionalitate gehitu zaizkio informazio honekin eroso aritzeko.



Irudia 9. Datuak bistartzeko GUI-aren diseinua

Proiektuan zehar jarraitu den hurrengo pausua, RR tarteen informazioa zehaztasunez behatzea izan da. Izan ere, balio hauek ezinbestekoak izan dira FA auresateko metodoa garatzeko kontuan hartu behar izan diren parametroen kalkuluak egin ahal izateko. Horregatik, RR tartek abiapuntutzat dituzten grafiko eta parametro ezberdinen emaitzak bistaratu ahal izateko, beste GUI baten erabilpena nahitaezkoa izan da. Horretarako, analisirako GUI tresna garatu da, 10. irudian erakusten dena.

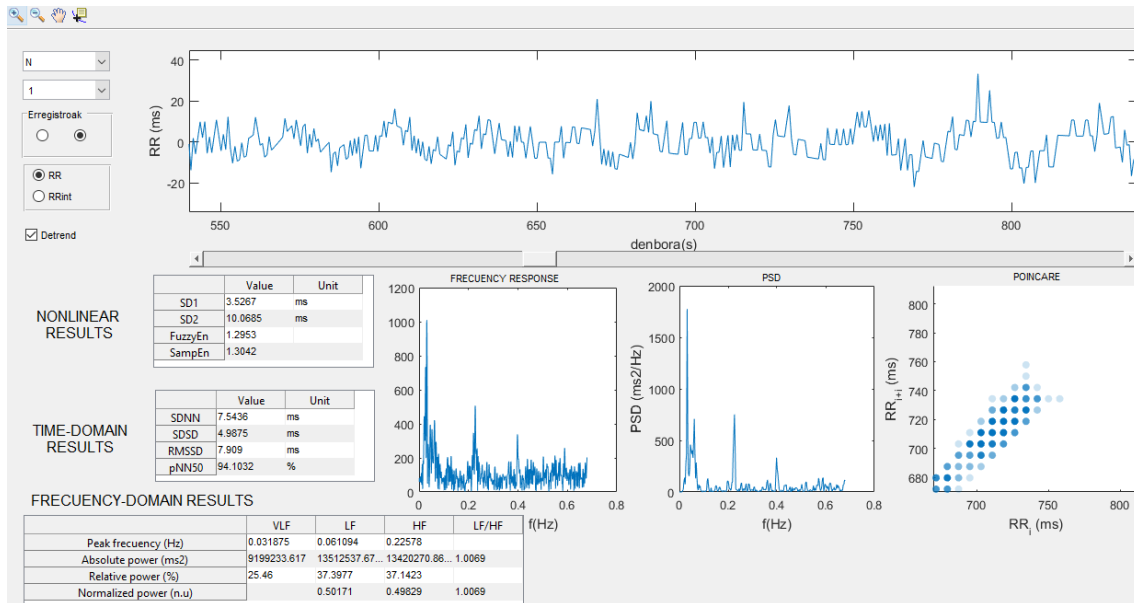
Bigarren GUI honek, beraz, RR tarteen grafikoa ere izango du, baina kasu honetan grafikoa bost minutuko tartean izandako balioekin kalkulatu da. Izan ere, parametroak bost minuturo kalkulatu egin behar dira. Honen bitartez ere, hiru grafika irudikatu ditugu seinalea aztertzeko. Batetik, RR tarteko seinaleari iragazkia aplikatu zaio seinalearen osagai jarraitua ezabatzeko, eta horren anplitude espektroa irudikatu da *'frequency response'* izeneko grafikan. Bestetik, *'Power Spectral Density'* izenburua duen grafikoa RR tarte jakin horren potentzia dentsitate espektrala (PDE) irudikatu da. Azkenik, hirugarren grafikoa *'Poincare'* deiturikoa, Poincare plot-a irudikatu da.

Grafiko horiek guztiaz gain, GUI honetan parametroen balioak agertzen dira, bai metodo ez-linealetik ateratakoak, zein maiztasun edota denbora domeinuan kalkulatuak, hauek guztiak tauletan adierazita daudelarik.

Bigarren GUI honek dituen funtzionalitateak hurrengoak dira:

- Aurreko kasuan bezala, bi menu irristakor izango dira paziente jakin baten informazioa bistaratu ahal izateko.
- Paziente bakoitzeko bi erregistro ditugunez, eta kasu honetan, RR tartek irudikatze grafiko bakarra dugunez, *'Erregistroak'* izeneko aukeran bi botoi bereizi dira, pazientearen lehenengo erregistroa, edo bigarrena irudikatu ahal izateko.
- RR tartekin lortutako denbora seriean laginen arteko distantzia ez da konstantea. Beraz, ondaz ondoko laginen distantzia uniformea izateko, RR tarteen seinalea interpolatu egin behar da. Horregatik, RR tarteen seinalea, edo RR tarteen seinale interpolatua irudikatze aukera izango dugu.
- RR tarteen seinaleari osagai jarraitua kendu nahi izanez gero, *'Detrend'* aukera dago.

- Seinalearen hurrengo bost minutuak bistaratu ahal izateko *slider-a*.



Irudia 10. Analisisirako GUI-aren diseinua

Behin hiru domeinutako parametro ezberdinen kalkuluak eginda izan eta gero, esanguratsuenak direnak aukeratuko dira. Izan ere, asko dira lortutako parametroak, eta soilik batzuk izango dira FA auresateko metodoa lortzeko erabiliko direnak.

6.3. PARAMETRIZAZIOA

Hortaz, parametroen kalkuluak burutu ondoren, hurrengo pausua hauen sailkapena egitea izango da. Izan ere, aurretik aipatu bezala, ez dira parametro guztiak erabiliko algoritmoa garatzerako orduan. Aukeratutako parametroak hurrengoak dira:

- Batetik, denbora domeinuko parametroak berezi dira. Kasu honetan, hauen inguruan aztertutako hiru parametroak erabiliko dira: SDNN, RMSDD eta pNN50, hain zuzen ere. Parametro hauen kalkulua RR tarteen balioei formula ezberdinak aplikatuz egin da.
- Bestetik, maiztasun domeinuko parametroak daude. Parametroen kalkuluak burutu ahal izateko, RR tartearen PDE kalkulatu behar da. Horren ondoren, espektroa hiru zatitan banatu behar da, maiztasun ezberdineko tartek berezi ahal izateko, - VLF, LF eta HF tartek, alegia-. Horregatik, banaketa hau kontuan harturik, asko dira domeinu honen barruan lortuko ditugun parametroak. Beraz, kasu honetan, parametro batzuk algoritmoaren kalkulua egiteko baztertuko dira. Hauek dira erabiliko ditugunak:
 - Maiztasun tarte bakoitzeko potentzia altuena duen maiztasunaren balioa, Hz-tan neurtuta, ingelesez '*Peak frequency*' deritzo.
 - Maiztasun tarte bakoitzeko potentzia absolutua, , ms^2 -tan neurtuta. Ingelesez '*Absolute power*' esaten zaio.
 - LF eta HF bandetan izandako potentzia kalkulatu da, honako hau unitate normalizatuetan dagoelarik. Ingelesez, '*Normalized power*' deritzo. Honen kalkulua burutzeko, tarte jakin horren potentzia absolutua, VLF bandaren potentzia absolutua eta potentzia absolutu totala hartu behar dira kontuan.

- LF eta HF banden arteko potentzia erlazioa kalkulatu da. Alde batetik, potentzia absolutuen balioak erabiliz, eta beste aldetik, potentzia normalizatuen balioak erabiliz.
- Azkenik, metodo ez-linealetatik lortutako parametroak aurki ditzakegu. Hauen barruan, hauek dira erabiliko direnak:
 - Hasteko, SD1 eta SD2 parametroak daude. Hauek biak, Poincare plot-ean elipsea sortzen duten aldean balioak direlarik.
 - Horrez gain, zenbait entropien balioak kalkulatu dira, SampEn eta FuzzyEn.

Beraz, behin erabiliko diren parametro guztiak bildu eta gero, hauen sailkatze gaitasuna aztertzea pasatuko da.

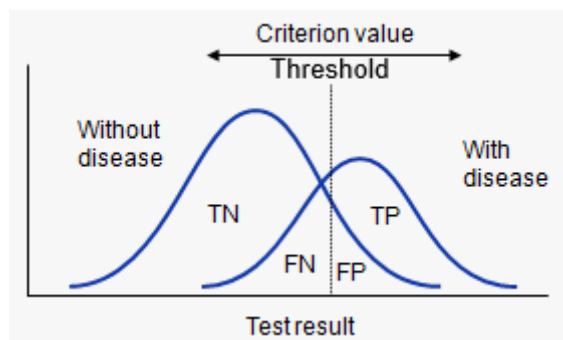
6.4. PARAMETROEN SAILKATZE GAITASUNA

Parametroen sailkatze gaitasuna aztertze hiru pausu jarraituko dira. Hasteko, parametro bakoitza indibidualki aztertuko da. Horren ondoren, parametro ezberdinen artean erlazioen bat dagoen ala ez ikertuko da. Azkenik, FA aurrerako metodoa lortzeko erabiliko diren sailkatzaile ezberdinak adieraziko dira.

6.4.1. PARAMETRO INDIBIDUALEN AZTERKETA

Parametro bakoitzaren azterketa, eta beraz, bere sailkatze gaitasuna aztertu ahal izateko, ROC kurba ikertuko da. Ingeleseko 'Receiver Operating Characteristic'-etik (ROC) dator, eta pazienteen ezaugarrietan oinarritutako aurreikuspen matematikoen bidez lortzen den grafikoa erakusten du. Laburbilduz, grafiko honek sistema sailkatzaile bitar baten diagnostikatze gaitasuna irudikatzen du, honen bereizketa ataria aldatzen doalako.

Grafikoa irudikatzeko, bi klase eduki dira kontuan. Batetik, paziente bakoitzetik izandako datuetatik lortutako parametro bat, eta bestetik, pazienteak FA duen ala ez bereizteko erabili den beste aldagai bat. Normalean, sailkapen bitarraren kasuan, klasearen aurreikuspena kasu bakoitzerako ausazko aldagai jarraitu batean, X , oinarritzen da. Balio horrekin, eta atalase jakin baten erabilpenaz (T , ingeleseko 'Threshold'), kasu horri positibotzat (P , $X > T$ betetzen denean) edo negatibotzat (N , $X < T$ betetzen denean) hartuko da. Horrez gain, alde bakoitzean, bi emaitza bereiz daitezke. Beraz, kasu bakoitzeko lau soluzio posible bereiz daitezke. Hurrengo irudian azaldutakoa ikus daiteke:



Irudia 11. Kasu bakoitza sailkatzeko irizpidea

Hortaz, sailkatzaile bitar batean aurki daitezkeen kasu posibleak hauek dira:

- Hasteko, benetako emaitza P baldin bada, eta aurreikuspen emaitza P izatea lortu bada, orduan benetako positiboa dela esan ohi da (*TP*, ingeleseko 'True positive').
- Alabaina, benetako emaitza N baldin bada, eta aurreikuspen emaitza P lortuz gero, orduan positibo faltsua dela esaten da (*FP*, ingeleseko 'False positive').
- Aitzitik, benetako emaitza N baldin bada eta aurreikuspen emaitza N izatea lortu bada, orduan benetako negatiboa sortu dela esan ohi da (*TN*, ingeleseko 'True negative').
- Azkenik, benetako emaitza P denean eta aurreikuspen emaitza N, orduan negatibo faltsua gertatuko da (*FN*, ingeleseko 'False negative').

Hau da, emaitza hauek gure proiektuan zera adierazten dute, alde batetik, aurreikusitako emaitzan pazienteen FA duela antzemanaz gero; hau da, positiboa. Baina berez, paziente hori osasuntsua baldin bada, orduan positibo faltsua (*FP*) den emaitza lortuko litzateke. Bestalde, aurreikusitako emaitza negatiboa baldin bada, honek pazienteen osasuntsu dagoela aditzera eramanez, eta benetan FA den gaixotasuna baldin bada, orduan, emaitza negatibo faltsua (*FN*) izango litzateke.

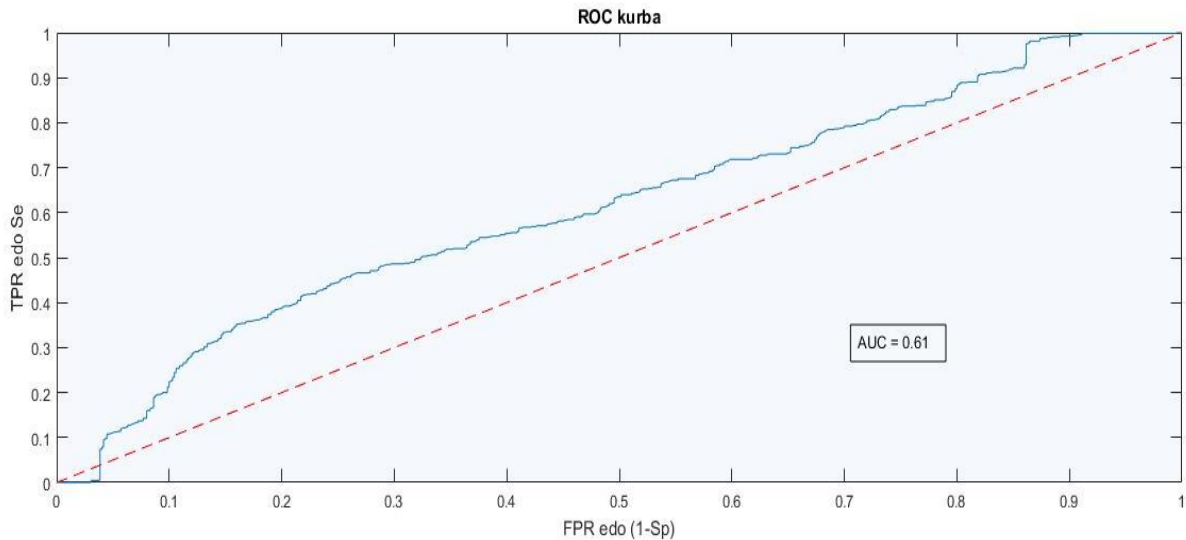
Horrez gain, lau emaitza hauekin oso adierazgarriak diren bi aldagai lor daitezke. Batetik, sentikortasuna (*Se*, ingeleseko 'Sensitivity') deituriko parametroa bereiz dezakegu. Bestetik, berriazkotasuna aurki daiteke (*Sp*, ingeleseko 'Specificity'). *Se* parametroak benetako positiboen tasa (*TPR*, ingeleseko 'True positive rate') adierazten du, *Sp* parametroak, berriz, benetako faltsuen tasa (*TNR*, ingeleseko 'True negative rate'), hurrengo formulak aplikatuz:

$$Se = \frac{TP}{TP + FN} \quad ; \quad Sp = \frac{TN}{TN + FP}$$

ROC kurba irudikatzerako orduan, bi aldagai hauek hartzen dira kontuan. Grafikoaren Y ardatzean, *TPR* irudikatzen da; hau da, *Se* aldagaiaren balioak, eta X ardatzean, ordea positibo faltsuen tasa (*FPR*, ingeleseko 'False positive rate') adierazten da; hots, $(1 - Sp)$ aldagaiaren balioak. 11. Irudian proiektuan erabilitako parametro jakin baten ROC kurba irudikatu da.

Grafika honetan, ROC kurbak 45 graduko zuzenaren itxura hartuko balu, horrek sentikortasunaren balioa berriazkotasunaren berdina dela adieraziko luke. Hau da, $Se = Sp$ atalasearen edozein baliorako. Kasu horretan, aztertutako parametroa ez luke informaziorik emango.

Erabilitako parametroaren sailkatze gaitasuna hobeto aztertzeko, ROC kurbaren azpiko area (AUC, ingeleseko ‘Area under ROC curve’) deituriko parametroa bereiz daiteke. AUC parametro honek sailkatzaileak ausazko kasu negatibo bat baino altuagoa den ausazko kasu positibo bat aukeratzeko probabilitatea adierazten du. 12. irudian, parametro jakin baten kasurako lortu den AUC -ren balioa aurkeztu da, ROC kurbaren eta 45 graduko zuzenaren artean aurki daitekeen aldea ere bereiziz.



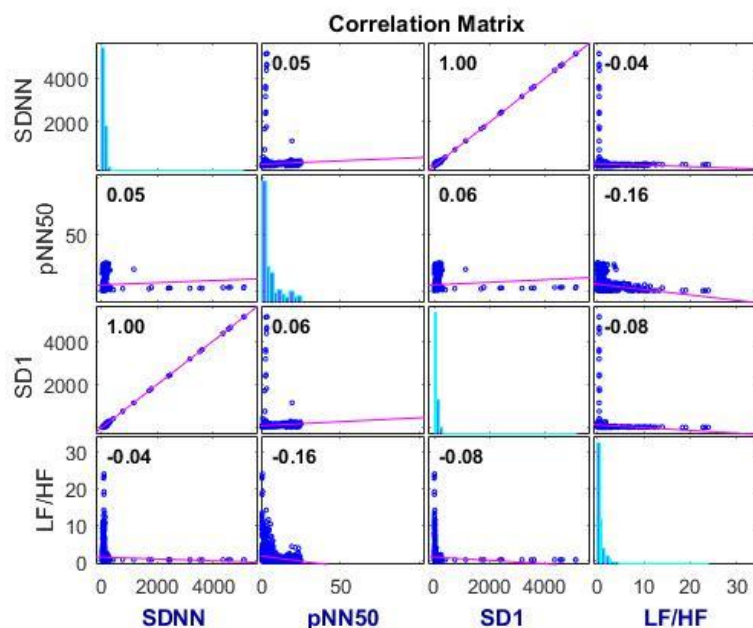
Irudia 12. Denborazko domeinuko parametro baten ROC kurba eta bere AUC balioa

6.4.2. PARAMETROEN ARTEKO ERLAZIOA

Orain, behin parametro bakoitzaren azterketa burutu eta gero, parametroen arteko erlazioa aztertu behar da. Horretarako, euren arteko korrelazioa ikertuko da. Hau da, atal honetan zehar aukeratutako parametro ezberdinak haien artean zerikusia duten ala ez aztertuko da.

Parametro bat beste batekin erlazionatzeko erabiltzen den parametroa ehunekotan adierazten da. Hortaz, parametro biren arteko korrelazioa neurri ezberdinetan aurki daiteke. Hasteko, parametro biren arteko korrelazio parametroa bat baliotik hurbil badago, horrek bi parametro horiek erabat erlazionatuta daudela adierazten du, eta beraz, algoritmoa garatzeko orduan bigarren parametro horrek ez luke informazio gehigarririk ekarriko. Bestetik, baliteke neurri batean korrelatuta egotea, baina ez guztiz. Azkenik, bata bestearekiko erlazionatzerik ez izatea gerta liteke.

Proiektu honetan korrelazio hauek antzeman ahal izateko, zenbait grafiko erabili dira. Horrez gain, korrelazio parametroen balioak kalkulatzeko beste funtzio erabili dira. 13. irudian erabilitako grafikaren adibidea ikus daiteke, non lortutako 4 parametroren arteko korrelazioa aztertu daiteke.



Irudia 13. Algoritmorako erabilitako lau parametroren arteko korrelazio matrizea

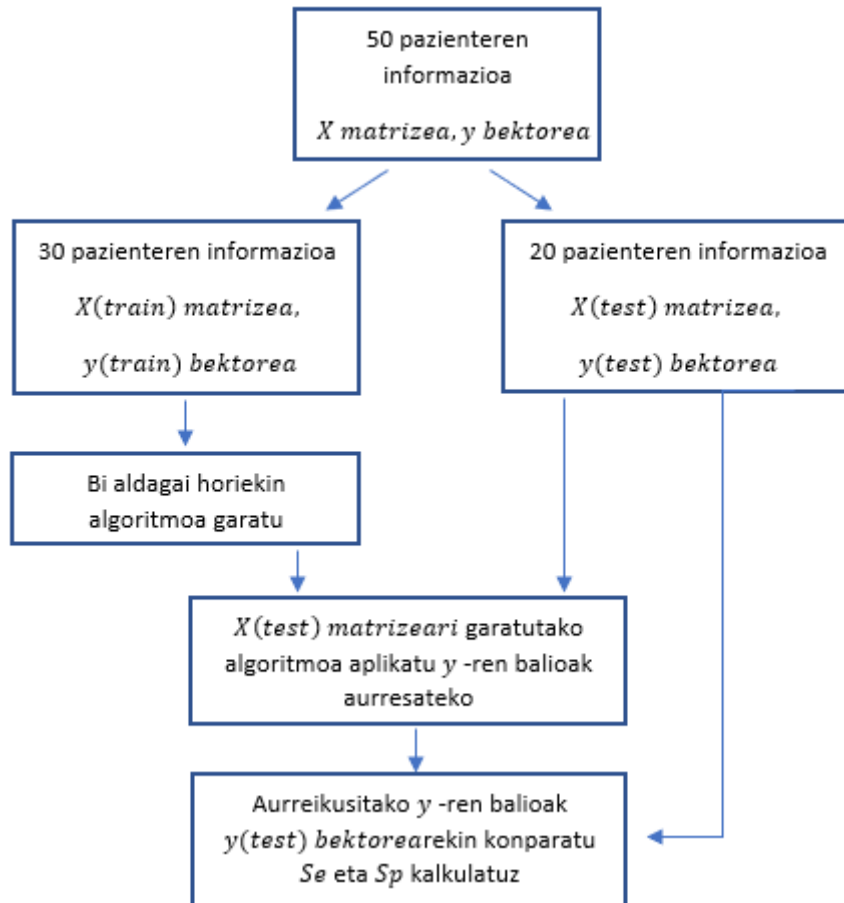
Irudi honetan, adibidez, SDNN eta SD1 parametroak erabat korrelatuta daudela antzeman daiteke. Izan ere, grafikoa 45 graduko zuzena irudikatu da, non balio guztiak zuzen horren gainean dauden, eta gainera, korrelazio parametroa 1.00 balioa duela antzematen da. Halaber, besteen artean erlaziorik ez dagoela esan daiteke, euren arteko korrelazio parametroa batetik oso urrun dauden balioak hartzen dituelarik.

6.4.3. SAILKATZAILEAK

Azkenik, parametroak indibidualki eta haien arteko erlazioak aztertu ondoren, FA aurreratzeko metodoa garatzeko ordua da. Horretarako, 50 pazienteei buruz izandako informazioa erabili da, horietako 25 paziente osasuntsuak direlarik, eta beste 25ak FA jasaten dutenak.

Algoritmoa burutu ahal izateko, bi izan dira erabilitako aldagaiak. Batetik, HRV analitika lortutako ezaugarriak gordeko dituen X matrizea. Bestetik, Faren benetako egoera gordetzeko erabiliko dugun y . Beraz, X matrizearen balio bakoitzari y -ren balio bat dagokio.

50 paziente horien banaketa gauzatu da, alde batetik, 30 pazienteren informazioa algoritmoaren garapenerako (Tr , ingeleseko 'train' moduan ezaguna dena) erabiliko da. Bestetik, gainontzeko 20 pazienteren informazioa test-erako (Te , ingeleseko 'test') erabiliko da. Hau da, algoritmoa $train$ barruan dagoen informazioarekin garatuko da. Horrela, algoritmoaren bitartez $test$ -eko pazienteek FA duten ala ez aurreikusiko da, eta azkenik, aurreikusitako balio hori, benetako balioarekin bat datorren ala ez antzeman egingo da. Horretarako, aurreko atalean azaldutako lau parametroen kalkulua burutuko da; hots, TP, FP, TN eta FN , eta hauekin, sentikortasuna (Se) eta berariazotasuna (Sp) aztertuko dira. Prozesua hobeto ulertzeko, 14. irudian horren eskema ikus daiteke.



Irudia 14. Algoritmoa garatzeko jarraitu den eskema

Paziente guztien informazioa bitan banatzeko prozesua 100 aldiz errepikatuko da, paziente bakoitzak garapen edota test barruan bereizteko ausazko funtzio bat erabiliko delarik. Horrela, metodo bakoitzeko Se eta Sp -ren balio ezberdinak lortu dira ehun kasu horietan. Azkenik, balio horien batezbestekoa eta desbideratze estandarra lortu dira, eta hauekin, erabilitako metodoa fidagarria den ala ez estatistikoki ondorioztatuko da, bootstrap deritzon teknika erabiliz.

Proiektu honetan hiru metodo ezberdin erabili dira FA auresateko, bakoitzarekin emaitza ezberdinak lortu direlarik. Hortaz, bakoitzaren oinarriko funtzionamendua azaltzea izango da hurrengo pausua.

6.4.3.1. MODELO LINEAL OROKORTUETAN OINARRITUTAKO SAILKATZAILEA

Ingeleseko 'Generalized linear models', erregresio linealaren familiako ereduen zabalkuntza dira. Eredu hauek hainbat eredu estatistikoak bateratzeko sortu zen; erregresio lineala, logistikoa eta Poissonena ere barnean sartuz.

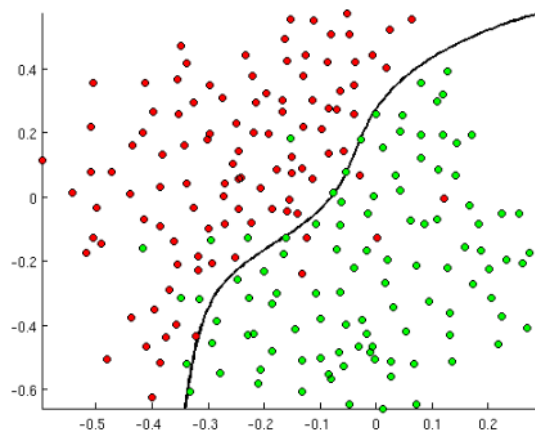
GLM-ak doiketa lineala orokortzen du eredu lineala erantzun aldagaiarekin erlazionatzea ahalbidetzen duelako. Hau lotura funtzio baten bidez gauzatu egiten da, baita ere neurketa bakoitzaren bariantza aurreikusitako balioaren funtzioa izatea onartzen delako.

Proiekturako erabili den lotura funtzioa binomiala izan da; hots, erregresio logistikoa. Hau da, aurreikusitako parametroa $[0,1]$ balioen artean egongo da. Beraz, benetako balioekin konparatu ahal izateko, aurreikusitako balio horiek 0 edo 1 baliora hurbildu behar dira. Horregatik, amaierako ereduak doiketa logistikoa moduan ezagutzen da.

6.4.3.2. EUSKARRI BEKTOREDUN MAKINAN OINARRITUTAKO SAILKATZAILEA

Ikasketa automatikoa (ingelesez 'Machine learning' moduan ezagunagoa), adimen artifizialaren adar bat da, zeinak adibide moduan emandako informazioan oinarrituz jokabide ezberdinak hedatzeko gai diren programak sortzea duen helburu. Honen barruan, euskarri bektoredun makinak (ingeleseko 'support vector machine', SVM) bereiz daitezke. Hauek ikasketa algoritmoekin erlazionatuta dauden gainbegiratutako ikasketa ereduak dira, sailkapenak egin ahal izateko hainbat datu aztertzea helburu dutelarik. Bi kategoria ezberdineko 'train' laginetan oinarrituz, SVM algoritmoak kategoria horietako bateko lagin berriei balio zehatzak esleituko dizkion eredu bat sortuko du, sailkatzaile lineal bitar ez probabilitistiko izango dena.

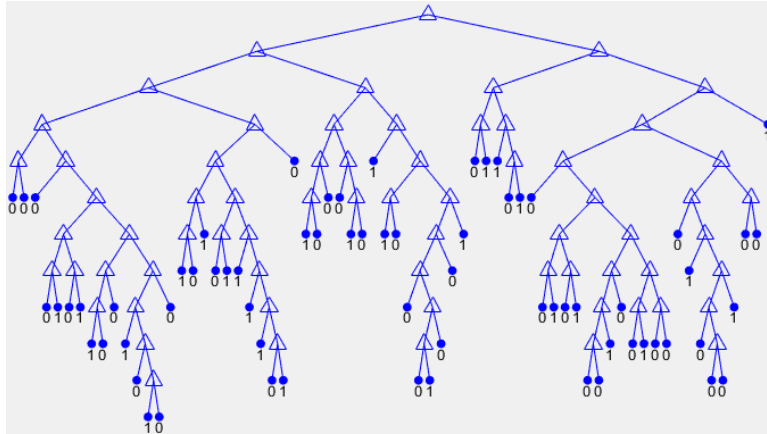
SMV ereduak espazioan kokatzen diren laginen irudikapena da. Espazio libre hori nahiko zabala izatea saiatuko da. Horrela, geroago lagin berriak espazio berean mapeatzerako orduan, zein kategoriaren barruan hartuko diren aurreikusiko da espazioko zein aldetan kokatu diren behatuz. 15. irudian SVM eredu bat ikus daiteke, non sailkapen kurba antzematen den.



Irudia 15. SVM ereduaren adibidea

6.4.3.3. ERABAKI-ZUHAITZEN BIDEZKO IKASKUNTZAN OINARRITUTAKO SAILKATZAILEA

Erabaki-zuhaitzen bidezko ikaskuntza (ingeleseko 'Decision tree learning'), aurreikuspen moduan erabaki-zuhaitza erabiltzen duen teknika da. Estatistikan, ikasketa automatikoan ere asko erabiltzen den ereduak da. Honen helburua, kasu berri baten aurrean bere klase-aldagaiaren balioa aurrerako ereduak sortzea da. 16. irudian proiekturako erabili den erabaki-zuhaitzaren adibidea ikus daiteke. Triangelu bakoitza aldagai iragarle bati egiten zaion test-a da, prozesua behin eta berriz errepikatuko delarik.



Irudia 16. Erabaki-zuhaitzen teknika

6.5. EMAITZEN LABURPENA

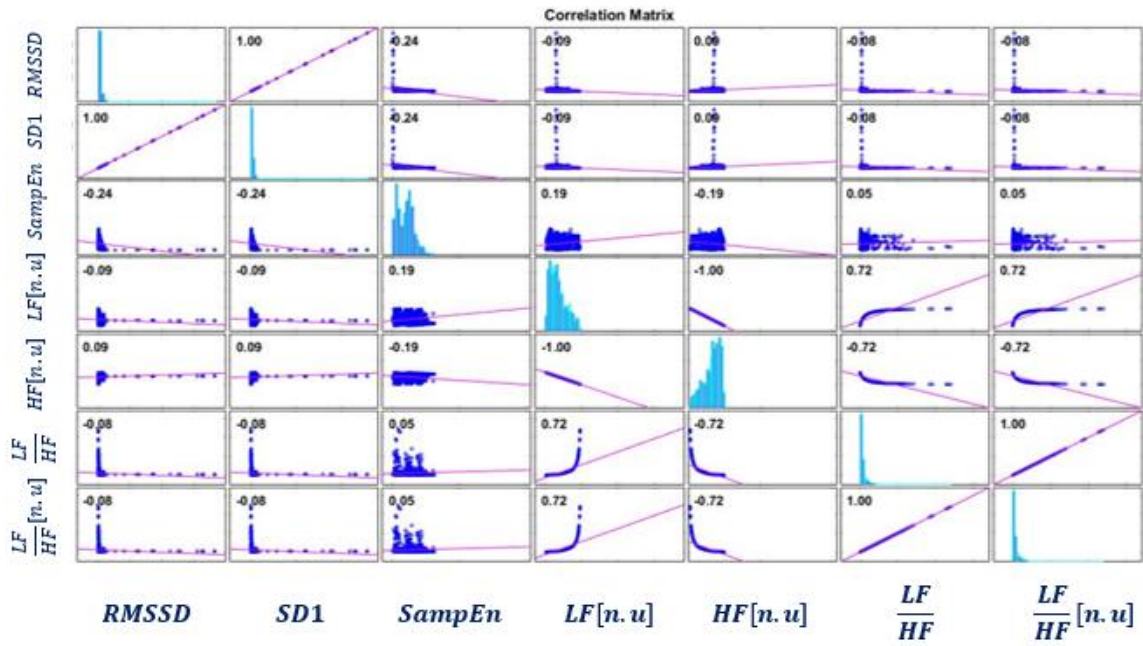
Proiektuan erabilitako sailkatzaile ezberdinak azaldu eta gero, atal honetan lortutako emaitzen laburpena azalduko da. Emaitzak modu zabalean II. Eranskinean deskribatzen dira.

Hasteko, parametro bakoitza indibidualki aztertu ondoren, parametro onenak sailkatuko dira. Horretarako, aurreko ataletan azaldu bezala, bakoitzaren ROC kurbari arreta jarriko diogu. Parametro guztiak aztertu ondoren, 7 izan dira lortu diren parametro hoberenak. Izan ere, euren AUC balioa [0.6, 0.75] balioen artean daude. Parametro horiek hurrengoak izan dira:

- Denbora-eremuaren barruan: $RMSSD$
- Metodo ez-linealen barruan: $SD1$ eta $SampEn$
- Maiztasun-eremuaren barruan: $LF[n. u]$, $HF[n. u]$, $\frac{LF}{HF}$ eta $\frac{LF}{HF}[n. u]$

Hala ere, AUC balioak tarte horien artean egoteak, egindako test-a hala-holakoa dela adierazten du. Horrez gain, parametro horien artean erlaziorik dagoen ala ez aztertu beharko da, zeren, haien artean erlaziorik egonez gero, horrek parametro gehigarriak daudela adieraziko luke, eta ondorioz, baztertu egin beharko lirake.

Hurrengo pausua, beraz, 7 parametro horien artean erlaziorik dagoen ala ez aztertzean lortu ditugun emaitzak azaltzea izango da. 17. irudian azkeneko parametro horien arteko korrelazio matrizea azaltzen da. Irudiari behatuz gero, $RMSSD$ eta $SD1$ parametroen artean erabateko korrelazioa dagoela antzeman daiteke. Horrez gain, maiztasun-domeinuko zenbait parametroen artean ere erabateko korrelazioa aurkitu da; $\frac{LF}{HF}$ eta $\frac{LF}{HF}[n. u]$ parametroen artean, eta $LF[n. u]$ eta $HF[n. u]$ parametroen artean, hain zuzen ere.



Irudia 17. 7 parametro onenen arteko erlazio azterketa

Azkenik, erabilitako hiru sailkatzaileekin lortutako emaitzak aztertu dira. Sailkatzaile hauekin lortutako Se eta Sp parametroen balioak aztertu eta gero, erabaki-zuhaitzen bidezko ikaskuntzan oinarritutako sailkatzailea emaitza onenak lortzen dituela ondorioztatu da. Sailkatzaile hau pazienteen ezaugarri parametroei aplikatuz sentikortasun eta berariazotasun parametroen batz bestekoa eta desbiderapena hurrengoak dira:

$$Se = 76.1 (8.33) \quad ; \quad Sp = 57.4 (13.81)$$

7. PLANGINTZA

Atal honetan zehar proiektuan burutu diren faseak adieraziko dira. Proiektua ataza ezberdinetan banatuko da. Hasteko, lan taldea deskribatuko da. Ondoren, fase guztien azalpena egingo da, eta azaldutakoa Gantt diagrama batean laburbilduko da. Azkenik, mugarriak deskribatuko dira.

7.1. LAN-TALDEA

Lehenengoko azpiatal honetan, proiektuaren garapenaz arduratutako lan-taldea era labur batean deskribatuko da, parte hartzaile bakoitzaren zereginak eta erantzukizunak ere definituz. Hurrengo taulan aipatutakoa era argi eta labur batean bistaratu da:

KODEA	ERANTZUKIZUNA	IZEN-ABIZENAK	ZEREGINA
K1	Proiektu zuzendaria	Unai Irusta Zarandona	Proiektua planteatzeaz, jarraitu beharreko pausuak adierazteaz eta, azkenik, dokumentua zuzentzeaz eta gainbegiratzeaz arduratzen dena.
K2	Junior ingeniaria	Itziar Eguskiza Garcia	Proiektua idazteaz eta garatzeaz arduratu dena. Hots, proiektuaren egilea.

Taula 1. Proiektuan parte hartzen duen lan-taldearen deskribapena

7.2. PROIEKTUA GARATZEKO GAUZATUTAKO FASEAK

Bigarren azpiatal honetan proiektua aurrera eramateko jarraitu diren fase ezberdinak definituko dira eginkizunak deskribatuz eta hauen iraupena zehaztuz.

7.2.1. PROIEKTUAREN KUDEAKETA

F 1-REN IRAUPENA: 252	HASIERA DATA: 2017/10/20	AMAIERA DATA: 2018/06/28
F 1.1	Proiektuaren proposamena	
Proiektu zuzendariak gradu bukaerako lanaren oinarriak azaldu junior ingeniariari.		
<ul style="list-style-type: none">• Hasiera data: 2017/10/20• Amaiera data: 2017/10/22• Baliabide teknikoak: Bibliografia (FA-ri buruzko artikulu eta liburuak).• Giza baliabideak: Proiektu zuzendaria (5 ordu).• Entregagaiak: Proiektuaren izenburua.		
M1: Proiektuaren hasiera. Proiektuaren proposamena onartu da. Horrez gain, proiektuan jarraituko diren pausuak finkatu dira.		
F 1.2	Proiektuaren kudeaketa eta jarraipena	

Proiektua behar bezala bideratzeko, eta hau egiten den heinean agertutako akatsen zuzenketak egiteko, proiektu zuzendariaren eta egilearen arteko bilerak.

- Hasiera data: 2017/10/23
- Amaiera data: 2018/06/28
- Baliabide teknikoak: Ordenagailua (30 ordu).
- Giza baliabideak: Proiektu zuzendaria (35 ordu) eta junior ingeniaria (10 ordu).

7.2.2. DOKUMENTAZIOAREN IRAKURKETA ETA MATLAB IKASTAROA

F 2-REN IRAUPENA: 54 HASIERA DATA: 2017/10/16 AMAIERA DATA: 2017/12/08

F 2.1 **MATLAB ikastaroa**

Ikerketa sail barruko kideek eskainitako MATLAB ikastaroa proiektuaren egileak programaren eta interfaze grafikoaren oinarriko ezagutzak bereganatzeko asmoarekin.

- Hasiera data: 2017/10/16
- Amaiera data: 2017/11/17
- Baliabide teknikoak: Bi ordenagailu (bat 20 ordu eta bestea 4) eta MATLAB lizentzia (bat 20 ordu eta bestea 4).
- Giza baliabideak: Junior ingeniaria (20 ordu) eta proiektu zuzendaria (4 ordu).

F 2.2 **Proiektuaren dokumentazioa**

Proiektuaren inguruan aurkitutako artikulu eta liburuen irakurketa.

- Hasiera data: 2017/10/24
- Amaiera data: 2017/12/08
- Baliabide teknikoak: Bibliografia (FA-ri buruzko liburu eta artikuluak) eta ordenagailu bat (20 ordu).
- Giza baliabideak: Junior ingeniaria (25 ordu).

M2: Proiektua definituta. MATLAB ikastaroa burutu, honen bidez eskuratutako ezagutzei esker, eta gaiari buruz informatu eta gero, erabiliko diren prozesu eta metodoak erabakiko dira. Beraz, proiektua definituta geratuko da.

7.2.3. BIHURKETEN GARAKETA

F 3-REN IRAUPENA: 2 HASIERA DATA: 2017/12/09 AMAIERA DATA: 2017/12/12

F 3.1 **Physionet-etik datuak jaitsi**

Physionet-eko web gunean eskuragarri izandako datu-basearen inguruko informazioa ordenagailuan gorde

- Hasiera data: 2017/12/09
- Amaiera data: 2017/12/10
- Baliabide teknikoak: Ordenagailu bat (ordu bat).
- Giza baliabideak: Junior ingeniari bat (ordu bat).

F 3.2 **.dat eta .qrs fitxategien bihurketa .mat formatura**

MATLAB programarekin lan egin ahal izateko, .dat eta .qrs fitxategiak .mat formatura bihurtuta.

- Hasiera data: 2017/12/11
- Amaiera data: 2017/12/12
- Baliabide teknikoak: Ordenagailu bat (3 ordu) eta MATLAB lizentzia (2 ordu).
- Giza baliabideak: Junior ingeniari bat (3 ordu).
- Entregagaiak: Bihurtutako .mat fitxategiak.

M3: Bihurketak burututa. Physionetetik hartutako fitxategiak .mat fitxategira ondo bihurtu dira.

7.2.4. DATU-BASEA GARATZEA

F 4-REN IRAUPENA: 65 HASIERA DATA: 2017/12/12 AMAIERA DATA: 2018/02/14

F 4.1 **Datu-basea sortu**

Proiektua garatzeko behar den informazio guztia era antolatuta batean gorde; hau da, datu-base batean gorde.

- Hasiera data: 2017/12/12
- Amaiera data: 2017/12/13
- Baliabide teknikoak: Ordenagailu bat (4 ordu) eta MATLAB lizentzia (4 ordu).
- Giza baliabideak: Junior ingeniari bat (4 ordu).
- Entregagaiak: Datu-basea.

F 4.2 **Fitxategiak bistaratzeko ingurunea garatu**

Fitxategiak bistaratzeko GUI-a sortu.

- Hasiera data: 2017/12/14
- Amaiera data: 2018/02/12
- Baliabide teknikoak: Ordenagailu bat (50 ordu) eta MATLAB lizentzia (50 ordu).
- Giza baliabideak: Junior ingeniari bat (50 ordu).
- Entregagaiak: Garatutako GUI-a

F 4.3 **Fitxategien azterketa**

Sortutako GUI-an dauden seinaleak aztertu.

- Hasiera data: 2018/02/12
- Amaiera data: 2018/02/14
- Baliabide teknikoak: Ordenagailu bat (3 ordu) eta MATLAB lizentzia (3 ordu).
- Giza baliabideak: Junior ingeniari bat (3 ordu) eta proiektu zuzendaria (2 ordu).

M4: Datu-basea definituta eta GUI-a garatuta. Ezinbestekoa den informazioa datu-basean bildu da eta informazio horrekin bistaratzeko tresna garatu da.

7.2.5. PARAMETROEN KALKULUEN BURUTZEA

F 5-REN IRAUPENA: 68 HASIERA DATA: 2018/02/19 AMAIERA DATA: 2018/04/27

F 5.1 **Parametroen balioak kalkulatu**

Parametroen kalkuluak burutzeko funtzioak garatu.

- Hasiera data: 2018/02/19

- Amaiera data: 2018/03/09
- Baliabide teknikoak: Ordenagailu bat (40 ordu) eta MATLAB lizentzia (40 ordu).
- Giza baliabideak: Junior ingeniari bat (40 ordu).
- Entregagaiak: Parametroen kalkuluak.

F 5.2 Parametroak aztertze ingurunea garatu

Kalkulatutako balioak aztertze GUI sortu, parametro horiek dagokien seinaleekin erlazionatuz.

- Hasiera data: 2018/02/26
- Amaiera data: 2018/04/25
- Baliabide teknikoak: Ordenagailu bat (50 ordu) eta MATLAB lizentzia (50 ordu).
- Giza baliabideak: Junior ingeniari bat (50 ordu).
- Entregagaiak: Garatutako GUI-a

F 5.3 Parametroen azterketa

Kalkulatutako parametroen balioak aztertu garatutako tresnaren bidez.

- Hasiera data: 2018/04/26
- Amaiera data: 2018/04/27
- Baliabide teknikoak: Bi ordenagailu (bat 2 ordu eta bestea ordu 2) eta MATLAB lizentzia (2 ordu bat eta 1 bestea).
- Giza baliabideak: Junior ingeniari bat (4 ordu) eta proiektu zuzendaria (3 ordu).

M5: Parametroen kalkuluak eta GUI-a garatuta. Parametroen kalkuluak burututa eta hauek bistaratzeko tresna garatuta.

7.2.6. ALGORITMOEN GARAPENA ETA EMAITZEN KALKULUA

F 6-REN IRAUPENA: 33 HASIERA DATA: 2018/04/30 AMAIERA DATA: 2018/06/01

F 6.1 Algoritmoen garapena

Kalkulatutako parametroen balioetatik abiatuta FA aurrerako algoritmoak garatu.

- Hasiera data: 2018/04/30
- Amaiera data: 2018/05/18
- Baliabide teknikoak: Ordenagailu bat (30 ordu) eta MATLAB lizentzia (30 ordu)..
- Giza baliabideak: Junior ingeniari bat (30 ordu).
- Entregagaiak: Algoritmoa.

F 6.2 Emaiza estatistikoak kalkulatu

Algoritmoa aplikatuz lortutako emaitza estatistikoak kalkulatu.

- Hasiera data: 2018/05/21
- Amaiera data: 2018/05/28
- Baliabide teknikoak: Ordenagailu bat (7 ordu) eta MATLAB lizentzia (6 ordu).
- Giza baliabideak: Junior ingeniari bat (7 ordu).
- Entregagaiak: Emaitzak.

F 6.3 Emaizen azterketa

Lortutako emaitzen azterketa burutu ondorioak atera ahal izateko.

- Hasiera data: 2018/05/28
- Amaiera data: 2018/06/01
- Baliabide teknikoak: Bi ordenagailu (bat 2 ordu eta bestea 2) eta MATLAB lizentzia (2 ordu bat eta ordu 1 bestea).

- Giza baliabideak: Junior ingeniari bat (4 ordu) eta proiektu zuzendaria (3 ordu).

M6: Algoritmoak garatuta eta emaitzak kalkulatuta. Proiektua burutzeko erabili diren algoritmoen garapena eta horiekin lortutako emaitza estatistikoen kalkulua burutu dira.

7.2.7. DOKUMENTAZIOA

LP 7-REN IRAUPENA: 29 HASIERA DATA: 2018/06/01 AMAIERA DATA: 2018/06/29

LP 7.1

Lana idatzi

Proiektuaren inguruko dokumentazioa idatzi da.

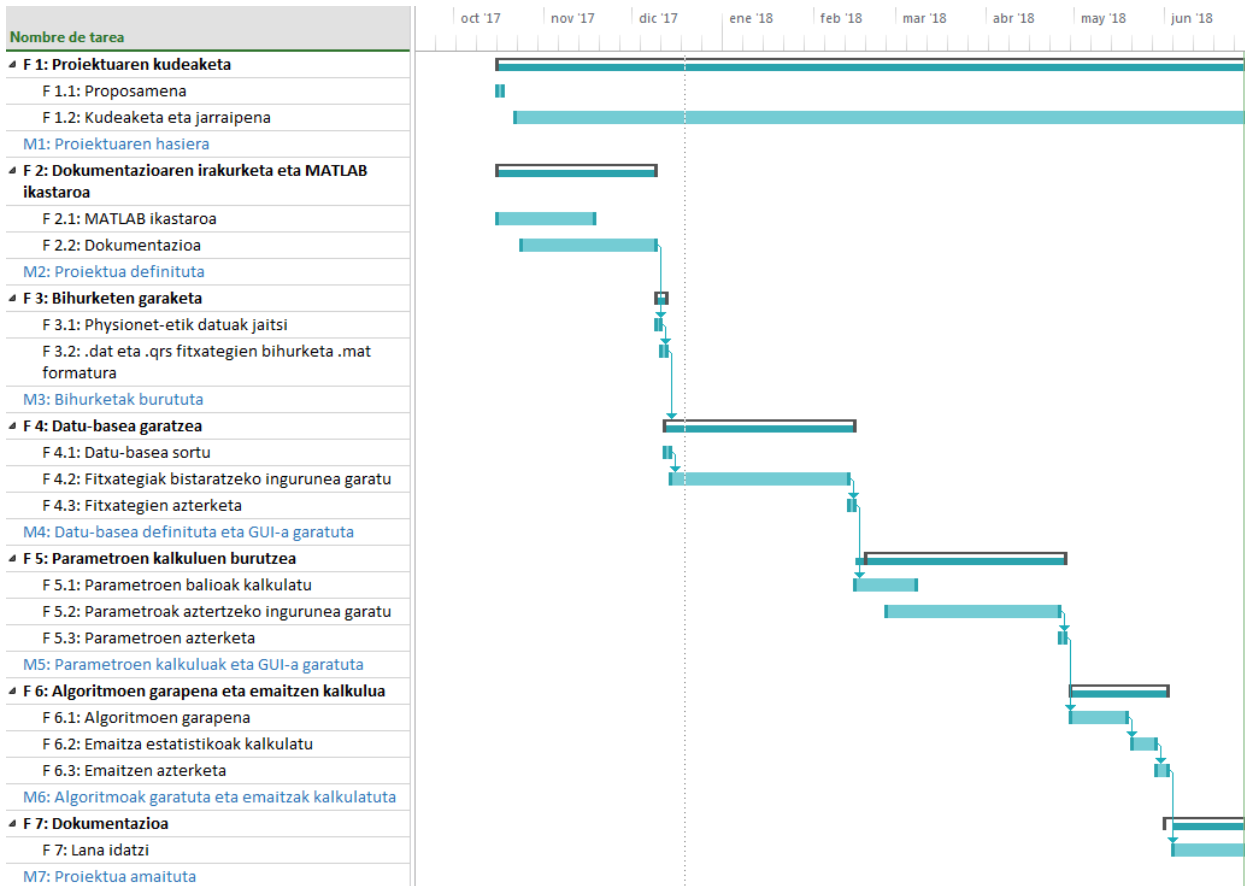
- Hasiera data: 2018/06/01
- Amaiera data: 2018/06/28
- Baliabide teknikoak: Ordenagailu bat (100 ordu).
- Giza baliabideak: Junior ingeniari bat (100 ordu).
- Entregagaiak: Memoria.

M7: Proiektua amaituta. Dokumentu guztia idatzi eta epe barruan entregatu da.

7.3. MUGARRIAK

MUGARRIAK	DESKRIBAPENA	DATA
M1	Proiektuaren hasiera	2017/10/22
M2	Proiektua definituta	2017/12/08
M3	Bihurketak burututa	2017/12/11
M4	Datu-basea definituta eta GUI-a garatuta	2018/02/14
M5	Parametroen kalkuluak eta GUI-a garatuta	2018/04/27
M6	Algoritmoak garatuta eta emaitzak kalkulatuta	2018/06/01
M7	Proiektua amaituta	2018/06/29

Taula 2. Mugarriak



Irudia 18. Gantt diagrama

8. AURREKONTUA

Atal honetan, proiektuan zehar izandako kostu totala azalduko da. Hortaz, aurrekontua burutzeko lau kostu eduki dira kontutan: barne orduak, amortizazioak, azpikontratazioak eta gastuak.

8.1. AURREKONTUAREN KONTU-SAILAK

8.1.1. BARNE ORDUAK

Azpiatal honetan proiektua aurrera eramateko eta burutzeko partaide bakoitzaren lanordu kopurua, eta horrek dakartzan kostuak azaltzen dira.

PARTAIDEA	KOSTUA (€/ORDUKO)	ORDU KOPURUA	PARTAIDEAREN KOSTUA (€)
Proiektu zuzendaria	60	52	3.120
Junior ingeniaria	30	350	10.500
Barne orduen kostu totala:			13.620

Taula 3. Barne orduen kostuak

8.1.2. AMORTIZAZIOAK

Azpiatal honetan proiektua garatzeko beharrezkoak izan diren baliabide materialei zein materialak ez direnei dagozkien kostuak azaltzen dira.

BALIABIDEA	KOSTUA (€)	BIZITZA ERABILGARRIA (ORDU)	ORDU KOPURUA	BALIABIDEAREN KOSTUA (€)
Proiektu zuzendariaren ordenagailua	800	2.000	38	15.2
Junior ingeniariaren ordenagailua	800	2.000	332	132.8
MATLAB 2016a bi lizentzia	2.000	1.700	215	253
Amortizazioen kostu totala:				401

Taula 4. Amortizazioen kostuak

8.1.3. AZPIKONTRATAZIOAK

Proiektu honetan zehar ez da azpikontrataziorik egin. Hortaz, ataza honen barruan ez da kosturik egongo.

8.1.4. GASTUAK

Azpiatal honetan proiekturako erabili diren eta proiektua bukatu ondoren berrerabili ezin izango diren baliabideen kostua azaltzen da. Hurrengo taulak hauen kostu totalen laburpena erakusten du:

BALIABIDEA	KOSTUA (€)
Bulego materiala	25
Argia	40
Kableak	3
16 GByteko pendive-a	12
Guztira	80

Taula 5. Gastuen laburpen taula

8.1.5. AURREKONTU OSOAREN LABURPENA

ATAZA	KOSTUA (€)
Barne orduak	13.620
Amortizazioak	401
Azpikontratazioak	0
Gastuak	80
Guztira	14.101

Taula 6. Guztizko kostu aitortpena

9. ARRISKUEN ANALISIA

Atal honen helburua proiektuaren garapenean zehar izan daitezkeen arriskuak adieraztea da. Proiektua amaituta dagoenez, arriskuak ekidin direla esan daiteke. Hala ere, hasieratik hauen aurreikuspena gauzatzea ezinbestekoa izan da, eta ondorioz, hauek ekar ditzaketen kalteen azterketa egin da.

Arrisku analisi hau aurrera eramateko bi kontzeptu hartu dira kontuan. Alde batetik, arriskuak berez gertatzeko duen probabilitatea aztertu da. Bestetik, arrisku horrek proiektuan izan dezakeen eragina ikertu egin da. Hortaz, arriskua balioesteko sasoian kontuan izango dira hau gertatzeko probabilitatea eta honek izan dezakeen eragina.

Aurreikusitako arrisku posibleak eta hauei aurre egiteko hartuko liratekeen kontingentzia neurriak hurrengoak dira.

9.1. KOSTUEN ARRISKUAK (A)

Aurreikusitako garapen kostuak gainditzea kostuen arriskuen barnean hartzen den kontzeptua da. Proiektu honen kasuan, hau posiblea izango litzateke kontuan hartu gabeko kostuak agertuko balira; hala nola, material gehiago beharrezkoa ziurtatzea, edota prezioetan aldaketak izatea. Hortaz, gertaera hauei aurre egiteko, aurrekontuan ezustekoak gehitu dira, gure kasuan, %5ekoak izanik.

Hala ere, gure proiektua garatu ahal izateko ez da material askorik erosi. Beraz, arrisku hau gertatzeko probabilitatea oso txikia da, eta proiektuan izango lukeen eragina ere txikia izango litzateke.

9.2. EGUTEGI ARAZOAK (B)

Gerta liteke proiektuan zehar garatutako lan planean zehaztutako datak gainditzea arrazoi ezberdinen ondorioz. Horrek; hau da, proiektuaren atalen batek atzerapenak izateak, proiektu osoa atzeratzea eragin dezake. Arrisku hau izateko probabilitatea ertaina da, eta bere eragina baxua.

Hau saihesteko, zehaztasunez neurtu beharko da proiektuaren atal bakoitza burutzeko beharko den denbora, eta epe muga hau gainditzea eragin ahalko luketen faktoreak aurreikusi.

9.3. DATUEN GALERA (C)

Proiektuan zehar erabiltzen diren datuak galtzea posiblea da. Datu hauek garatutako tresnak edota datu-basea izan daitezke, adibidez. Arrisku hau gertatzeak ondorio desatseginak ekarriko lituzke. Izan ere, lan berdina berriro egin beharko litzateke, eta horrek, atal horri eskainitako denbora gainditzea eragingo luke.

Hala ere, arrisku hau gertatzeko probabilitatea baxua da, datuak dispositibo ezberdinetan gorde eginez gero, arriskua saihestuko litzatekelako. Baina, arriskua ematekotan, honen eragina oso altua izango litzateke.

9.4. ARRISKU TEKNIKOAK (D)

Proiektuan zehar garatutako tresnetan arazoak izatea posiblea da. Arazo hauen adibidea izango litzateke garatutako kodean akatsen bat izatea, eta horrek, behar bezala ez funtzionatzea eragingo luke. Horrez gain, arrisku hauen artean interfazearen diseinuarekin erlazionatutako arazoak ere multzo honetan sartuko lirateke.

Arrisku hau agertzeko probabilitatea ertaina da, eta eragina handia. Beraz, hau saihesteko bi neurri hartu dira. Alde batetik, garatuko diren interfazeak hasiera batetik oinarrizko kalitatea beteko dute, era berean, erabiltzailearentzat horietan zehar aritzeko atsegina egingo zaio. Bestalde, kodean ager daitezkeen akatsei dagokienez, garatutako tresna bakoitza zatika-zatika probatuko da kodearen atal bakoitza ondo dagoela ziurtatzeko.

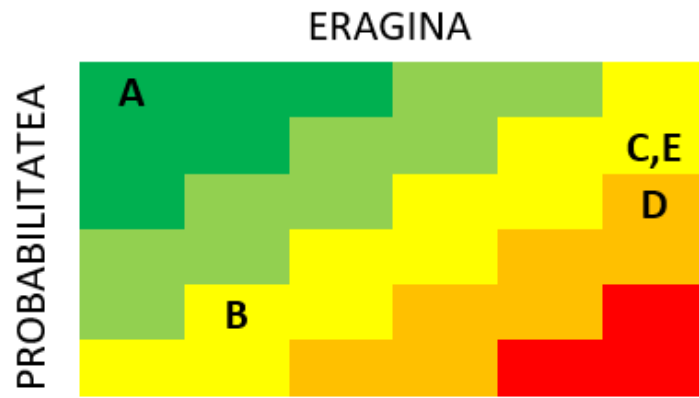
9.5. ARRISKU TEKNOLOGIKOAK (E)

Proiektua garatzeko ezinbestekoak diren teknologien erabilpenean ezagutzak ez izatea, hardware/software arazoak izatea, edota garatutako interfaze edo script fitxategien integrazioan arazoak izatea dira multzo honen barruan ager daitezkeen arriskuen adibideak.

Hala ere, arrisku hauek saihesteko zenbait neurri hartu dira. Hasteko, GUI interfazearen erabilpena lehendabizikoa izateagatik, proiektua garatzen hasi aurretik honi buruzko ikastaroa egin da. Horrez gain, erabiliko diren ekipoekin arazoak ez izateko, hauek behar bezala erabiltzen eta mantentzen saiatuko da. Hortaz, esandakoa kontutan hartuz, arrisku teknologikoak agertzeko probabilitatea baxua da eta eragina altua.

9.6. ARRISKU ANALISIAREN LABURPENA

Jarraian, aipatutako arrisku ezberdinen probabilitate-eragina erlazioa erakusten duen matrizea azaltzen da.



Irudia 19. Proiektuaren arriskuekin erlazionatutako probabilitate-eragin matrizea

10. ONDORIOAK

Proiektu honekin, FA auresateko gai den metodoa garatzea saiatu da. Horretarako, hainbat pausu jarraitu dira.

Hasteko, FAP duten pazienteen eta paziente osasuntsuen datuak lortu dira, eta informazio hori era dinamiko eta eroso batean landu ahal izateko zenbait tresna garatu dira. Horren ondoren, informazio horretatik abiatuta, HRV analisirako esanguratsuak diren hainbat parametroen kalkuluak burutu dira. Horrez gain, parametro horiek hobeto aztertu ahal izateko, beste ingurune bat garatu da.

Hori guztia burutu ondoren, kalkulaturako parametro horietan oinarritzen diren hiru metodo garatu dira, bakoitzak algoritmo hori bideratzeko ezaugarri ezberdinak erabiltzen dituelarik.

Azkenik, garatutako algoritmo horiek fidagarriak diren ala ez antzemateko zenbait parametroen kalkuluak burutu dira. Aipaturako azken parametro hauek sentikortasun eta berariazkotasun parametroak direlarik.

Parametro bi hauek, gaixotasun bat auresaterako orduan ager daitezkeen lau kasuetan oinarritzen dira. Hau da, aurretik aipaturako *TP*, *FP*, *TN* eta *FN* parametroetan, hain zuzen. Hortaz, datu berriei algoritmoa aplikatuz, eta lortutako balio horiek benetako balioekin konparatzean, lau parametro horien balioak lortu dira. Horren ondoren, sentikortasun eta berariazkotasun parametroen batezbesteko eta desbideratze balioak kalkulatuak direlarik.

Proiektu honetan, erabilitako hiru algoritmoetatik, erabaki-zuhaitzen bidezko ikaskuntzan oinarrituta dagoen metodoak balio esanguratsuenak lortu ditu. 390 kasu aztertu direla kontuan hartuz, hauek dira lau parametro horien inguruan lortutako batezbestekoak:

- *TP* barruan sailkatu diren 148 kasu egon dira. Hau da, gaixotasuna zeukaten 148 paziente aurkitu dira.
- *TN* barruan 121 kasu eman dira; hots, gaixotasuna ez zuten, eta osasuntsu bezala sailkatu diren 121 emaitza lortu dira.
- *FP* barruan sailkatutako 84 kasu egon dira. Honek pazienteak gaixotasuna duela antzeman dela adierazten du, benetan horrela ez denean.
- *FN* barruan 48 kasu eman dira. Azken honek, pazienteak gaixotasuna ez duela aurreikusi dela adierazten du, benetan baiezko kasua denean.

Emaitza hauek aztertuta, metodoa oso fidagarria ez dela ondoriozta daiteke. Egia da benetako kasuak kasu faltsuak baino askoz gehiago direla, baina, hala ere, faltsuak diren hainbat kasu antzeman dira.

Azkenik, parametro horiek formula barruan hartzen dituzten parametroen balioak kalkulatu dira; hots, sentikortasuna eta berariazkotasuna, non hauen emaitzetan ondorioztutakoa ere antzeman daitekeen. Hona hemen azken parametro hauen batezbestekoa eta desbideratze balioak, hurrenez hurren:

$$Se = 76.1 (8.33) \quad ; \quad Sp = 57.4 (13.81)$$

11. ERREFERENTZIAK

- [1] Gaixotasun kardiobaskularrengatiko heriotza tasa. IMQ Corporation.
- [2] When the Beat is off-Atrial Fibrillation. Publications of American Stroke Association.
- [3] Wolf CD, Rudd AG. The Burden of Stroke White paper: Raising awareness of the global toll of stroke-related disability and death.
- [4] World Health Organisation. The atlas of heart disease and stroke.
- [5] Avoiding heart attacks and strokes: don't be a victim – protect yourself. Publications of the World Health Organisation, 2005.
- [6] Rathore SS, Berger AK, Weinfurt KP, Schulman KA, Oetgen WJ, Gersh BJ, et ál. Acute myocardial infarction complicated by atrial fibrillation in the elderly: prevalence and outcomes. *Circulation*. 2000; 101(9):969-74.
- [7] Pathophysiological determinants of worse stroke outcome in atrial fibrillation, *Cerebrovascular Disease* 2010;30(4):389-95.
- [8] Amar D, Roistacher N, Zhang H, Baum MS, Ginsburg I, Steinberg JS. Signal-averaged P-wave duration does not predict atrial fibrillation after thoracic surgery. *Anesthesiology* 1999;91:16-23.
- [9] Kannel WB, Abbott RD, Savage DD, McNamara PM. Epidemiology of chronic atrial fibrillation: the Framingham Study. *N Engl J Med* 1982;306:1018-22.
- [10] Sudlow M, Rodgers H, Kenny RA, Thomson R. Population based study of use of anticoagulants among patients with atrial fibrillation in the community. *Br Med J* 1997; 314:1529–30.
- [11] Petersen P, Godfredsen J. Embolic complications in paroxysmal atrial fibrillation. *Stroke* 1986; 17:622–5.
- [12] Godtfredsen J. Atrial fibrillation, etiology, course and prognosis: a follow-up of 1212 cases. In: Kulbertus HE, Olsson SB, Schlepper M, eds. *Atrial Fibrillation*. Molndal, AB Hassle, 1982:134–45.
- [13] Kannel WB, Abbott RD, Savage DD. Coronary heart disease and atrial fibrillation: the Framingham Study. *Am Heart J* 1983; 106:389–96.
- [14] Al-Khatib SM, Wilkinson WE, Sanders LL, McCarthy EA, Pritchett EL. Observations on the transition from intermittent to permanent atrial fibrillation. *Am Heart J*. 2000;140(1):142-145.
- [15] I. Daskalov and I. Christov. Improvement of resolution in measurement of electrocardiogram RR intervals by interpolation. *Med Eng Phys*.1997; 19(4):375–379.
- [16] M. Brennan, M. Palaniswami, and P.Kamen. Do existin measures of Poincaré plot geometry reflect nonlinear features of heart rate variability. *IEEE Trans Biomed Eng*, 48(11): 1342-1347,2001.
- [17] GB Moody, AL Goldberger, S McClennen, SP Swiryn. Predicting the Onset of Paroxysmal Atrial Fibrillation: The Computers in Cardiology Challenge 2001. *Computers in Cardiology* 28:113-116 (2001).

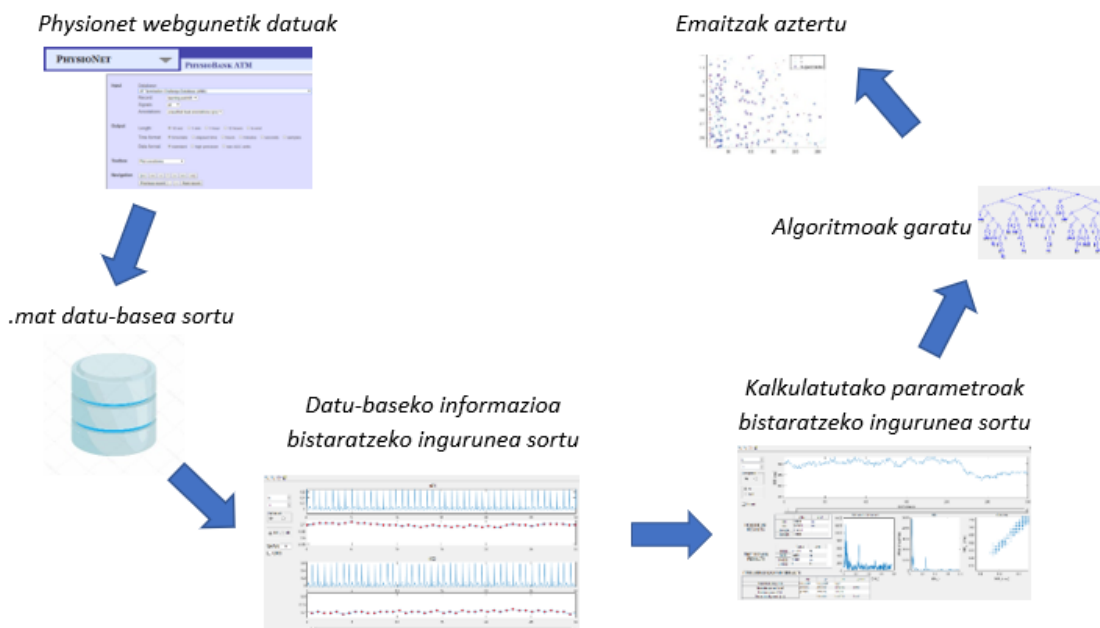
[18] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220.

ERANSKINAK:

- I. Eranskina: Erabiltzailearen gida
- II. Eranskina: Emaidza zabalagoak

ERANSKINAK

Proiektua aurrera eraman ahal izateko jarraitu diren pausuak 20. irudian azaltzen dira, non garatutako bi inguruneak agertzen diren, horren ondoren algoritmoak emandako emaitzen azterketa egin dela adierazi da, beste zenbait atalez gain. Atal honek bi zati bereiziko ditu, alde batetik, I. Eranskinean sortu diren bi inguruneak azalduko dira, eta bestetik, II. Eranskinean lortutako emaitzak azalduko dira.



Irudia 20. Proiektua garatzeko jarraitu den eskema

I. ERANSKINA: ERABILTZAILEAREN GIDA

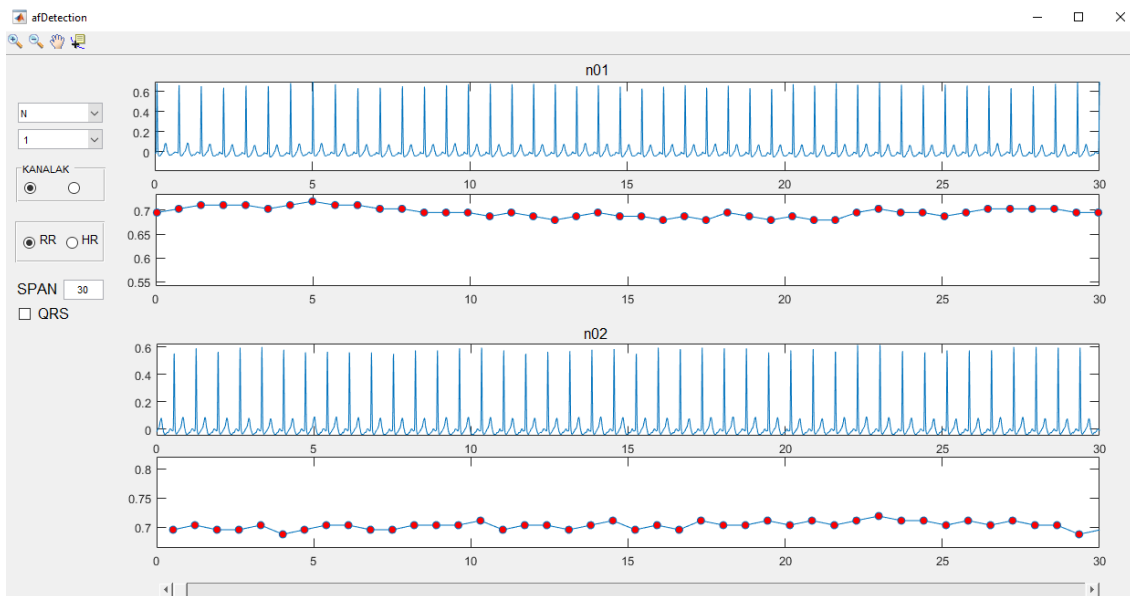
Atal honetan beraz, proiektuan zehar erabili diren bi inguruneen deskribapena egingo da, honako hauek zertarako garatu diren eta zer funtzionalitate duten adieraziz.

DATU-BASEKO INFORMAZIOA BISTARATZEKO INGURUNEA

Behin proiekturako erabili behar diren datuak datu-base batean behar bezala egituratuta izan eta gero, hurrengo pausua, datu horiek era dinamiko eta eroso batean bistaratzeko aukera eskainiko duen tresna garatzea da. Horretarako, *afDetection* izena duen ingurunea sortu da.

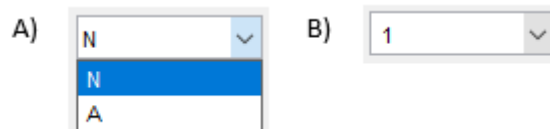
Tresna hau abiarazteko, MATLAB programan *afDetection* jarri beharko da, horren ondoren *enter* botoiari sakatuz. Horrela, aipatutako ingurunea abiaraziko da 21. irudian agertzen den leihoa bistaratzuz. Irudiarri begiratzuz gero, hasieratik lehenetsitako seinaleak bistaritzen direla

antzeman daiteke. Izan ere, ingurunea kargatzean lehenengo pazientearen erregistroak bistaratzeko nahia izan da.



Irudia 21. afDetection hasierako leihoa

Ingurunea hobeto ulertzeko, bertan agertzen diren aukerak azalduko dira. Hasteko, paziente jakin baten informazioa bistaratzeko bi menu irristakor erabili dira, 22. irudian ikus daitezkeen moduan. Batetik, paziente osasuntsu edo FAP duen pazienteak aukeratzeko aukera dago, 21A irudian ikus daitezkeen moduan, non 'N' sakatuz gero pazienteak osasuntsuen artean bereiziko diren, eta 'A' aukeratzuz gero, paziente gaixoen artean. Bestetik, bigarren aukera, paziente multzo horien barnean paziente bat edo beste aukeratzeko da bere identifikazio zenbakia aukeratzuz, 22B irudian ikus daitezkeenez.

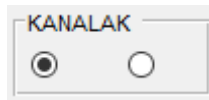


Irudia 22. Menu irristakorrak

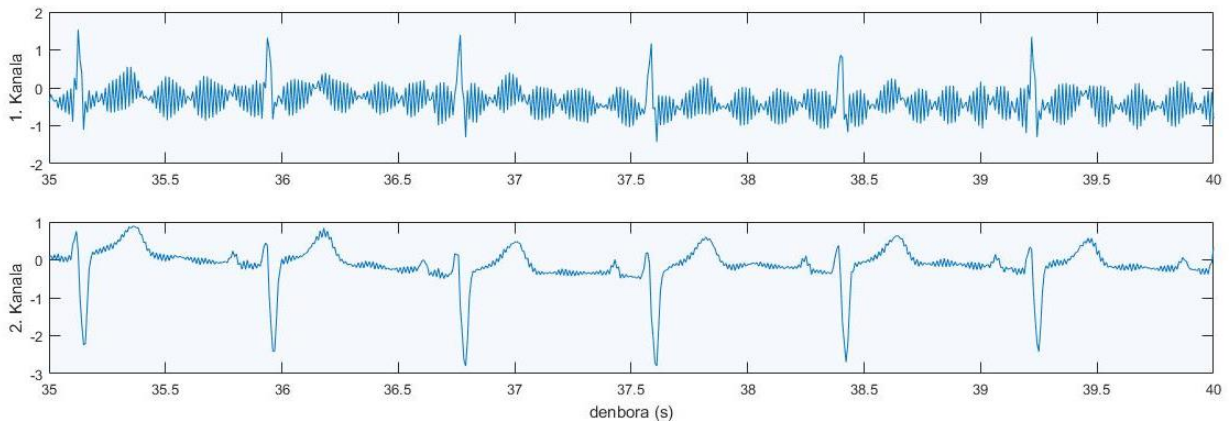
Behin pazientearen aukeraketa egin eta gero, bere informazioarekin lortu ditugun seinaleei erreparatu diegu. Paziente bakoitzeko bi erregistro ezberdin bereiz daitezke, aurretik aipatutako 21. irudian ikus daitezkeen moduan, adibide horretan paziente jakin horren erregistroak 'n01' eta 'n02' dira hurrenez hurren, grafiken goialdean azaltzen den bezala. Gainera, erregistro bakoitzeko bi seinale ezberdin antzeman daitezke:

- Batetik, EKG seinalea dago. Aurreko ataletan aipatu den bezala, paziente bakoitzeko gordetako EKG seinalea bi kanaletan bereizita dago, kanal bakoitzak gorputzaren gune ezberdin batetik lortutako informazioa duelarik. Bi kanalak grafiko berean gainjarri ez daitezkeen, kanal bat edo beste bistaratzeko aukera gehitu da, 23. irudian ikusten denez. Horrela, momentu jakin batean kanal batetik zarata gehiegi antzematen bada, beste kanaletik informazio adierazgarriagoa lortzeko aukera ematen da. 24. irudian honen

adibidea ikus daiteke, non paziente bereko EKG seinalearen une berean bigarren kanaletik seinale garbiagoa lortzen den.



Irudia 23. EKG seinalearen kanal ezberdinen aukeraketa

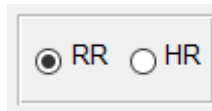


Irudia 24. Bost segundoko tarte batean antzemandako bi kanalen adibidea

- Bestetik, RR tarteak aurkezten dituen grafikoa bereiz daiteke, EKG seinale bakoitzaren azpialdean. Grafikoaren puntu gorri bakoitzak QRS konplexuaren barruan R uhina gertatzen den unea adierazten du. Geroago, puntu gorri horiek guztiak lotzen dituzten zuzenak marraztu dira. RR tarteek, ondoz ondoko taupaden arteko denbora adierazten dute. Horregatik, grafiko horren emaitzak 0.7 segundo inguruko balioak hartzen ditu. Hala ere, RR tarte horietatik, HR deituriko tarteak (ingeleseko 'Heart Rate') lor daitezke, zeintzuk minutuko izan den taupada kopurua adierazten duten. HR tarteak hurrengo formula aplikatuz lortzen dira:

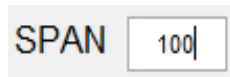
$$HR(i) = \frac{60}{RR(i)}$$

Hortaz, formulari erreparatuz gero, RR tarteak eta HR tarteak elkarrekin erlazionatuta daude. Alde batetik, RR tarteek taupaden arteko denbora tarte adierazten dutelarik, eta bestetik, HR tarteek minutuko zenbat taupada izan diren adierazten dute. Horregatik, erregistro bakoitzaren bigarren grafika honetan, RR tarteak edota HR tarteak bistartzeko aukera izango da, 25. irudian ikus daitekeen moduan botoi bata ala bestea sakatuz.



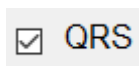
Irudia 25. RR edo HR tarteak bistaratzeko aukera

Horrez gain, grafikoen denbora tarteak aldatzeko aukera gehitu da. Horrela, bi seinaleen iraupena 30 segundotik 120 segundora bitartean alda daiteke. Hortaz, grafikoaren denbora tarteak aldatzeko, 'SPAN' aukeraren eskuz sartu beharko da nahi den balio berria, 26. irudian ikusten den moduan.



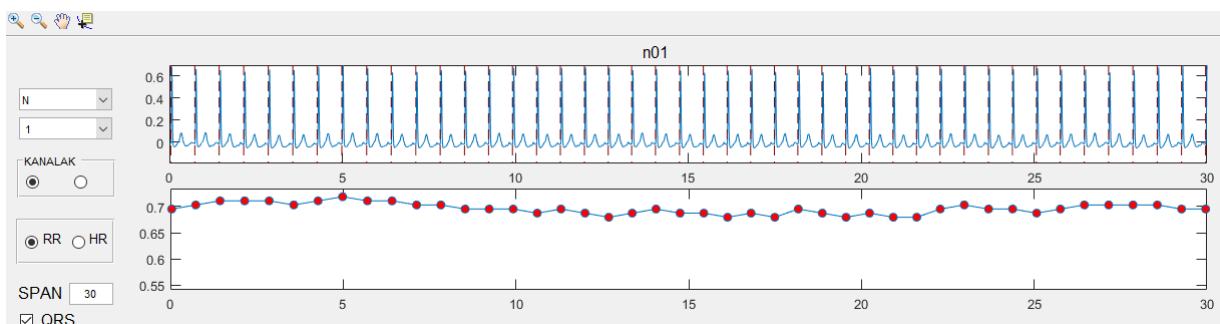
Irudia 26. Seinaleak ehun segundoko iraupenarekin bistaratzearen adibidea

Halere, ingurune honek QRS konplexuaren barruan R uhina gertatzen den uneak bistaratzeko aukera eskaintzen du, 'QRS' aukeraren klikatuz, 27. irudian ikus daitezkeen moduan.



Irudia 27. R uhinen aldiuneak adierazteko aukera

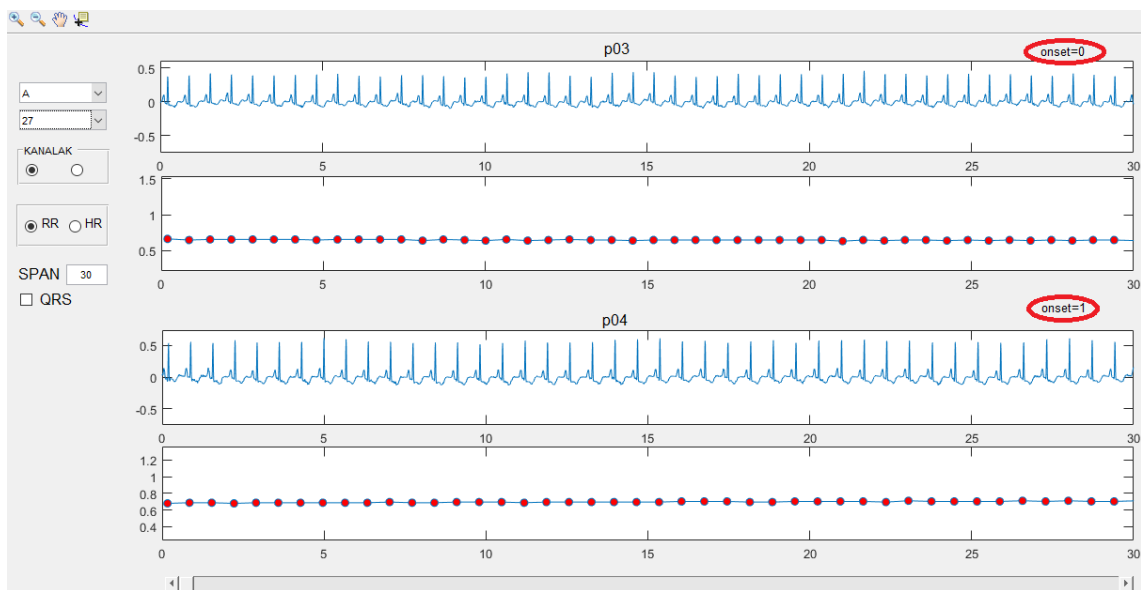
28. irudian, aipatutako azken aukera hau sakatuz lortuko litzatekeen emaitza adierazten da, non EKG seinalean R uhina gertatutako uneetan marra bertikal gorriak antzeman daitezkeen. Kasu honetan, pazientearen erregistro bakarra bistaratu delarik.



Irudia 28. Paziente jakin baten lehenengo erregistroa R uhinen aldiuneekin

Gainera, seinaleak aurrera eta atzera eramateko, -slider erramintaren bidez-, baita seinaleen denbora zati batzuk handiagotzeko edota txikiagotzeko, zein puntu jakin batean markadore bat jartzeko aukerak gehitu dira, - 28. irudiko ezker aldeko goiko aldean ikus daitezkeen ikonoak erabiliz-.

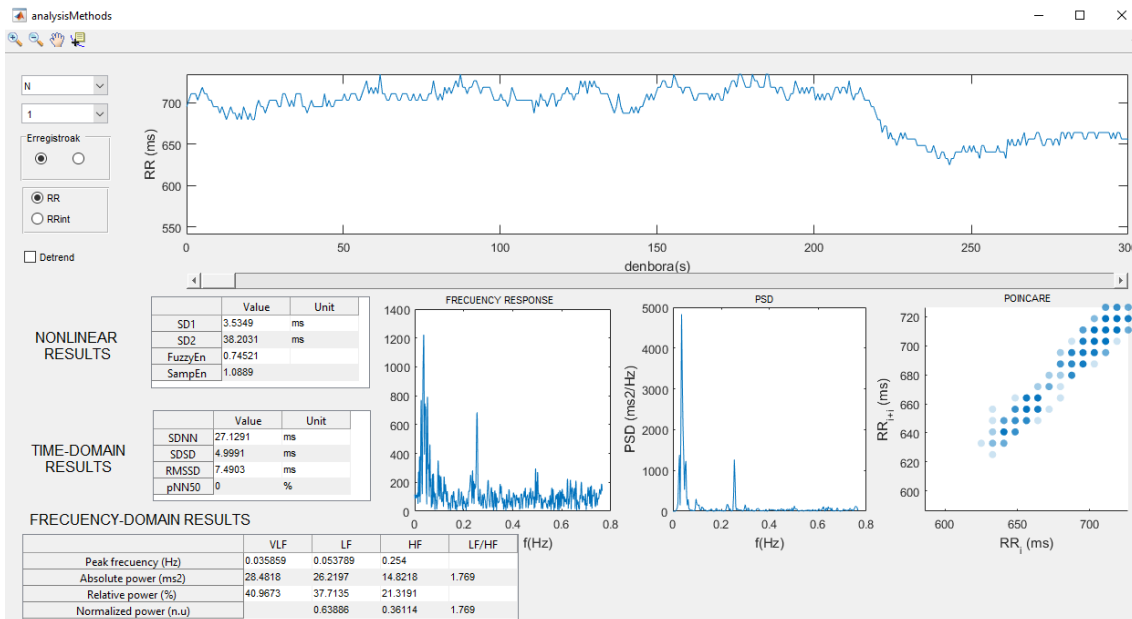
Azkenik, tresna honen erabileraren inguruan, pazienteen buruzko ezaugarriei begira, bada paziente osasuntsu eta FAP duten pazienteen artean ezberdintasun bat. 'N' motakoak diren pazienteak ez dute arritmia, eta hortaz, paziente bakoitzeko dauden erregistro biren artean ez dago ezberdintasun handirik, soilik minutu, ordu edota egun ezberdinetan hartutako informazioa dutela. Bestalde, 'A' pazienteak FA dute, eta horrek bi erregistroak une esanguratsuetan hartu direla adierazten du, erregistro bat bestearengandik desberdintzera behartzen duena. Hortaz, 29. irudian ikus daitekeen moduan, bi erregistroak bereizi dira. Izan ere, arritmia duten paziente bakoitzaren erregistro bakoitiak,- adibide honetan 'p03' izena duena-, paziente hori FAP gertaera izan ez duen unean lortu dela adierazten du, eta horregatik 'onset = 0' moduan adierazi da. Horren kontrara, paziente bakoitzaren erregistro bikoitiak,- adibide honetan 'p04' dena-, FAP gertatu baino bost minutu lehenagoko datuak gordetzen ditu; hots, 'onset = 1'. Honako hau bereiztea oso inportantea da.



Irudia 29. FAP duten pazienteen erregistroen arteko desberdintasuna

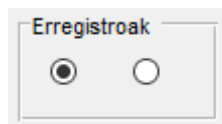
ALGORITMOA GARATZEKO KALKULATU DIREN PARAMETROAK BISTARATZEKO INGURUNEA

Garatutako bigarren tresna honek, aurrekoak zituen zenbait funtzionalitate ere aurkezten ditu, baina, horrez gain, algoritmoa garatzeko erabilitako parametroen ezaugarriak ere adierazten ditu. Kasu honetan, ingurunearen izena *analysisMethods* da eta 30. irudian, honek duen itxura ikus daiteke.



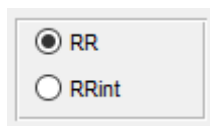
Irudia 30. analysisMethods hasierako leihoa

Bigarren tresna hau, algoritmoak garatzeko erabili diren parametroak aztertzeko erabili da. Algoritmoak garatzeko abiapuntutzat hartu diren balioak RR tarteenak izan direnez, kasu honetan, paziente bakoitzaren inguruan seinale hau soilik bistaratu da. Hori bai, lehenengo ingurunean bezala, bi menu irristakor erabiliko dira paziente bakoitzaren informazioa bistartzeko aukeratu ahal izateko. Grafiko bakarra irudikatu denez, eta paziente bakoitzeko bi erregistro bereizten direnez, erregistro bat edo beste sailkatzeko aukera ematen da, 31. irudian ikusten den moduan.



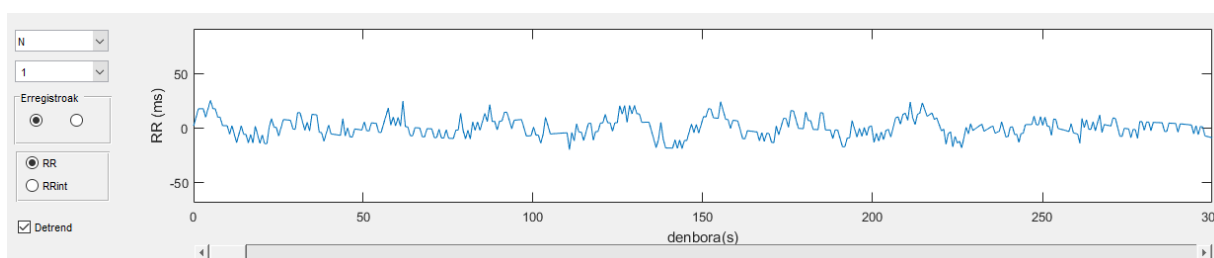
Irudia 31. Paziente bakoitzeko bistaratu nahi den erregistroaren aukeraketa

RR tartek adierazten dituen seinaleetan laginen arteko distantzia ez da konstantea. Hala ere, maiztasun-domeinuko parametroak kalkulatu ahal izateko erabili behar den seinalearen laginen arteko distantziak distantziakidea izan behar du. Horregatik, RR tarteei interpolazioa deituriko prozesua aplikatu behar zaie. Horrela, lagin berriak sortuko dira 0.25 segunduro (laginketa maiztasuna 4 Hz). Beraz, ingurune honetan, RR tartek dituen seinalea edota RR tarte interpolatuak dituen seinalea bistartzeko aukera gehitu da, 32. irudian ikus daitekeen moduan botoi bat edo beste sakatuz.



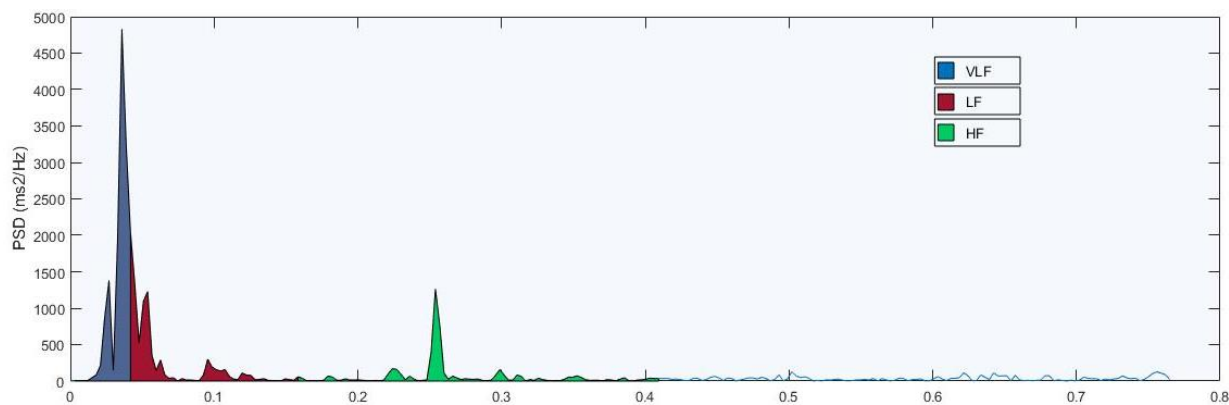
Irudia 32. RR seinalea edo RR seinale interpolatua bistaratzeko aukera

Horrez gain, aurretik azaldu den moduan RR tarteak ondoz ondoko taupaden artean igarotako denboraldiak gordetzen dituela jakina da. Seinale honetan garrantzitsuena dena, ordea, seinaleko laginen artean dagoen bereiztasuna aztertzea da, eta honako hau, daukagun seinaleari batezbesteko osagaia kenduz ere lor daiteke. Aukera hori bera tresna honetan ematen da 33. irudian ikusten den moduan 'Detrend' deituriko aukeran klikatuz. Horrela, aurreko 30. irudian ez bezala, non RR tarteak balioak 700 milisegundotik hurbil dauden, kasu honetan zeroren inguruan aurkitzen dira, eta azaldu den moduan, bata bestea besteak beste informazio bera erakusten du. Izan ere, seinale hauetan garrantzia duena horien arteko aldaketa da.



Irudia 33. RR seinaleari batezbestekoa kenduta (detrend) lortzen den emaitzaren adibidea

Azkenik, ingurune honetan zehar antzeman daitezkeen gainontzeko atalak parametro ezberdinen kalkuluak dira. Horretarako, domeinu bakoitzerako taula bat sortu da, non eremu horren barnean dauden parametroen izenak azaldu dira kasu bakoitzean lortutako emaitza ondoan adieraziz. Horrez gain, parametroen balioak hobeto aztertzeko zenbait grafiko sortu dira. Horrela, kasu bakoitzean 'Poincare plot'-ean sortzen den elipse itxura antzeman daiteke, adibidez, edota 'PSD' grafikan maiztasun tarte bakoitzeko potentzia altuena duen maiztasunaren balioa ikus daitezke, 34. irudian ikus daitezkeen moduan, non maiztasun banda bakoitza argi eta garbi bereiz daitezkeen.



Irudia 34. PSD grafikoaren adibidea, maiztasun tarte ezberdinak bereiziz

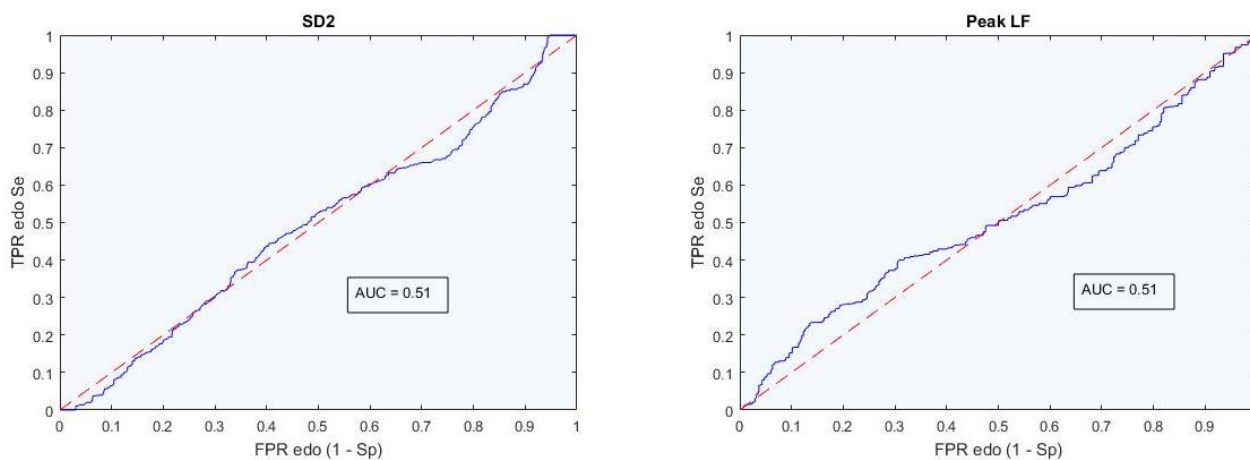
II. EMAITZA ZABALAGOAK

Bigarren eranskin honetan, parametroak kalkulatu eta gero balio horiekin lortutako emaitza ezberdinak azalduko dira. Horretarako, parametro indibidualen azterketaren bidez parametroen sailkatze gaitasuna nola egin den azalduko da. Geroago, parametro horien konbinaketa eginez euren arteko erlazioak aztertzeko erabilitako erramintak aztertuko dira. Azkenik, erabilitako algoritmo ezberdinen emaitzak azalduko dira.

PARAMETRO INDIBIDUALEN AZTERKETA

Parametro indibidualen azterketa egiteko, aurretik azaldu den moduan, ROC deituriko kurba irudikatu da funtzio jakin baten bitartez. Horrela, parametro bakoitzaren sailkatze gaitasuna ondorioztatu ahal izan da.

Hori egin eta gero, parametro batzuk algoritmoa garatzerako orduan ezer gutxirako balio dutela ohartu gara. Izan ere parametro horiei ROC kurba lortzeko funtzioa aplikatuz gero, AUC parametroaren balioa ere lortzen da, eta kasu hauetan azken parametro honen balioa %50-etik oso hurbil dagoela aztertu da. 35. irudian ikus daitekeen moduan, *SD2* eta *Peak LF* parametroak dira aipatutako parametroak.

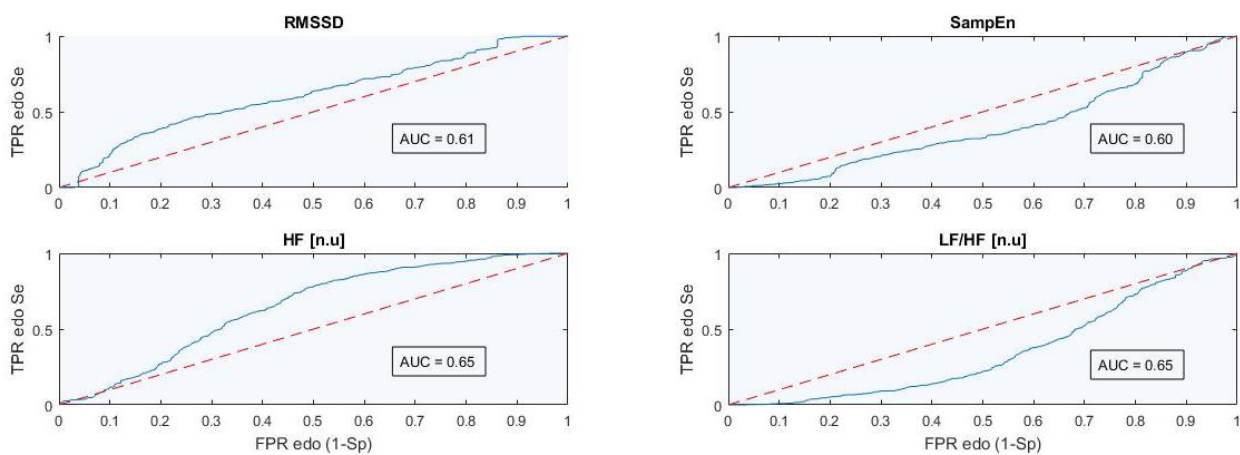


Irudia 35. *SD2* eta *Peak LF* parametroen ROC kurbak eta AUC balioak

Halaber, informazio esanguratsua eskaini ahal duten parametroak landu dira, horien artean, AUC balioa $[0.6, 0.75]$ tartearen barne daukaten zazpi parametro aurkitu dira, eta 0.6 baliotik oso gertu dauden beste bi parametro. 7. taulan parametro horiek aurkeztuko dira, euren AUC balioa ere bistaratuz, eta 36. irudian parametro horien artean bereizitutako lauren ROC kurba ikus daiteke.

<i>Denbora-domeinuko parametroak</i>		
	<i>RMSSD</i>	<i>AUC = 0.61</i>
<i>Maiztasun-domeinuko parametroak</i>		
Potentzia absolutuak	<i>HF(ms2)</i>	<i>AUC = 0.58</i>
	<i>LF[n.u]</i>	<i>AUC = 0.65</i>
Potentzia normalizatuak	<i>HF[n.u]</i>	<i>AUC = 0.65</i>
	$\frac{LF}{HF}$	<i>AUC = 0.65</i>
Erlazioak	$\frac{LF}{HF} [n.u]$	<i>AUC = 0.65</i>
<i>Parametro ez-linealak</i>		
Poincare	<i>SD1</i>	<i>AUC = 0.61</i>
	<i>SampEn</i>	<i>AUC = 0.60</i>
Entropiak	<i>FuzzyEn</i>	<i>AUC = 0.59</i>

Taula 7. Parametro hobereenak eta bere AUC balioa

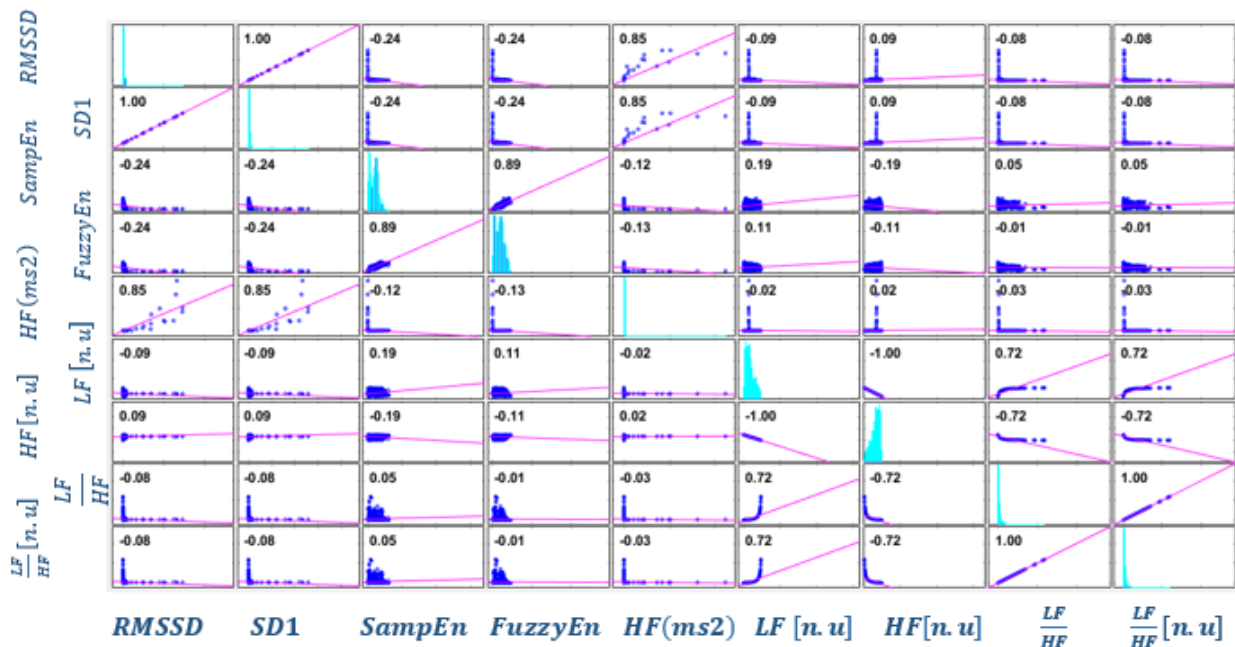


Irudia 36. Lau parametro onen ROC kurba eta haien AUC balioa

PARAMETROEN ARTEKO ERLAZIOAREN AZTERKETA

Orain, parametro hoberen moduan sailkatu ditugun artean haien artean korrelaziorik existitzen den ala ez aztertzeke erabili diren metodoak azalduko dira. Horretarako, aurretik esan bezala, korrelazio matrizea sortuko duen funtzioa erabiliko da.

Matrize hori sortzeko, 7. taulan azaldu diren parametroak aukeratu dira. Izan ere, horiek dira informazio gehien ematen duten parametroak, baina haien artean erlazioa existitzen bada ez dute informazio gehigarririk eskaintzen. 37. irudian parametro horien arteko korrelazio matrizea aurkezten da.



Irudia 37. Parametro hoberen korrelazio matrizea

Hortaz, matrizean lortutako balioak landuz, zenbait ondorio atera dira. Batetik, badaude erabateko korrelazioa duten parametro bikoteak. Bestalde, erabatekoa ez den erlazioa baina korrelazio koefiziente oso altua duten parametroak bereiztu dira. Azkenik, oso erlazio gutxi edota erlazioirik ez duten parametroak aurkitu dira. 8. taulan erlazioatuta dauden parametro bikoteak aurkeztu dira, eta hauen arteko korrelazio koefizientearen balioa, R batez adierazi dena, balio absolutuan.

Erabateko erlazioa duten parametroak		
<i>RMSSD</i>	eta	<i>SD1</i>
$\frac{LF}{HF}$	eta	$\frac{LF}{HF}[n.u]$
<i>HF [n.u]</i>	eta	<i>LF[n.u]</i>
<i>R = 1.00</i>		
Erlazio handia duten parametroak		
<i>RMSSD</i>	eta	<i>HF(ms2)</i>
<i>SD1</i>	eta	<i>HF(ms2)</i>
<i>SampEn</i>	eta	<i>FuzzyEn</i>
<i>HF [n.u]</i>	eta	$\frac{LF}{HF}$
<i>HF[n.u]</i>	eta	$\frac{LF}{HF}[n.u]$
<i>R = 0.85</i>		
<i>R = 0.85</i>		
<i>R = 0.89</i>		
<i>R = 0.72</i>		
<i>R = 0.72</i>		

Taula 8. Korrelatutako parametro bikoteak eta haien korrelazio koefizientea

GARATUTAKO ALGORITMOEK SORTUTAKO EMAITZAK

Azkenik, kalkulaturako parametroekin garatutako algoritmoen inguruan lortu diren emaitzak aztertuko dira. Horretarako, aurretik aipatu den moduan, sentikortasun eta berariazkotasun parametroen batezbestekoa eta desbiderapena aztertu dira.

Proiektuan zehar hiru algoritmo garatu dira, eta bakoitzarekin emaitza ezberdinak lortu dira. Emaitzak aztertu ondoren, soluzio onenak lortzen dituen metodoa erabaki-zuhaitzen bidezko ikaskuntzan oinarrituta dagoena dela antzeman da *Se* eta *Sp* parametroen batezbestekoa eta desbideratzea hurrengoak izanik:

$$Se = 76.1 (8.33) \quad ; \quad Sp = 57.4 (13.81)$$

Gainontzeko beste bi metodoen artean balio horietara gehien hurbiltzen zaion algoritmoa modelo lineal orokortuetan oinarritzen den sailkatzailea da. Honekin lortutako emaitzak hurrengoak izanik:

$$Se = 66.2 (14.11) \quad ; \quad Sp = 53.3 (16.73)$$

Azkenik, emaitza kaskarragoak lortu dituen metodoa euskarri bektoredun makinan oinarritutako sailkatzailea izan da:

$$S_e = 56.0 (48.16) \quad ; \quad S_p = 47.0 (46.22)$$