

# Greining á flogaköstum í heilaritum

## Gervigreind

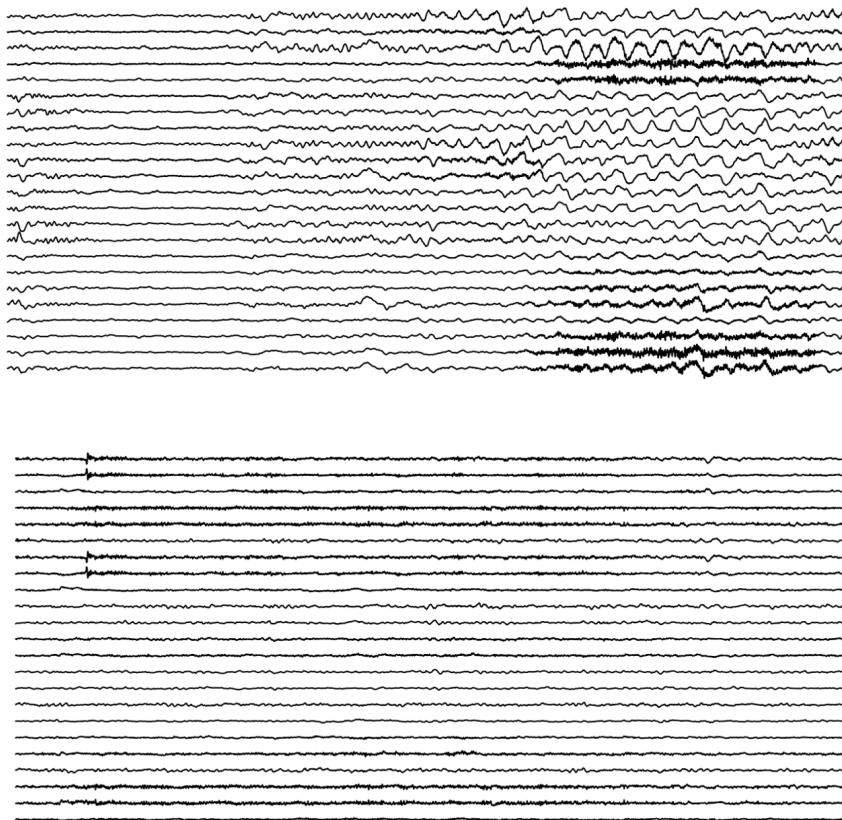
Alexander Guðjónsson

Emil Gauti Friðriksson

Apríl 2019

## 1 Inngangur

Viðfangsefni þessarar skýrslu er greining EEG heilarita og þá sérstaklega hvernig greina má mun heilarita einstaklinga með flog og án floga. Mynd 1 sýnir dæmi um heilbrigt heilarit og flogakast. Fundin voru ýmis auðkenni gagnanna og út frá þeim voru þjálfaðir og fínstilltir þrír flokkarar: Stoðvigravél, lógistísk aðhvarfsgreining og AdaBoost. Flokkurunum var síðan beitt í einstaklingsmiðuðu mati, þá er þjálfaður flokkari á gögnum allra sjúklinganna nema þanns einstaklings sem flokkarinn er síðan prófaður á. Ásamt flokaranna þriggja þá var einnig smíðað tauganet út frá rófríti EEG mælinganna.



Mynd 1: Heilarit af flogi (uppi) og heilbrigt heilarit (niðri)

Sjálfvirk greining EEG heilarita hvort það sýni flog eða ekki gæti orðið verðmætt verkfæri lækna og sjúklinga þeirra sem þjást af flogum. Enn í dag þurfa sérfræðingar oft að rýna í heilarit handvirkt til að greina hvort um hafi verið að ræða flog eða ekki. Þetta er kjörið verkefni fyrir vélnám og gervigreind því magn gagna sem safnast er oft mjög mikið.

## 2 Aðferðir

### 2.1 Gögn

Gögn notuð í þessu verkefni voru fengin úr doktorsverkefni Ali Shoeb[5] úr MIT. Gögnin samanstanda af 14 sekúnda heilaritum flokkuð eftir hvort þau sýni flog eða ekki. Safnað hefur verið saman 170 flogamælingum og 370 mælingum án floga, samtals 540 mælingum. Hver þessara 540 mælinga inniheldur svo 23 tímaraðir. Mælingar eru teknar með svokölluðu 10-20 kerfi[2].

### 2.2 Auðkenni

Þegar litið er á tímaraðirnar sjálfar sem inntaksgögn þá eru þau alls ekki lýsandi enda liggur áhuginn ekki á einstökum gildum. Í tímaröðum sem lýsa heilaritum liggur áhuginn frekar á að skoða heildarmyndina út frá eiginleikum sem byggja á tíðni raðanna, kallaðir rófseiginleikar (e. spectral features). Í okkar greiningu reiknuðum við aflrófspéttleika (e. power spectral density) fyrir mismunandi tíðnibönd (kölluð  $\alpha, \beta, \delta, \theta$ ) og Hjorth gildi hverrar tímaraðar og notum þær stærðir frekar sem inntök í flokkarana okkar [4].

Notast var við aðferð Welch's við þá reikninga sem má finna í `signal` pakkanum hjá `scipy`. Einn galli við þessar stærðir er að þær týna upplýsingum um tíma í röðunum en í heilaritum má gjarnan finna forboða fyrir flogaköstin svo við viljum geta nýtt upplýsingar um tíma í greiningu okkar. Við gerðum það með því að skipta tímaröðunum, sem eru 14 sekúndur niður í minni tímabil og skoða rófseiginleikana á hverju bili fyrir sig. Við fundum bestu stærðina fyrir tímarammann með því að þjálfa flokkarana okkar á öllum mögulegum stærðum og bera saman prófunarnákvæmni þeirra.

### 2.3 Flokkarar

Við bárum saman þrjár gerðir flokkara, lógistíska aðhvarfsgreiningu, stoðvigravél og safnaðferðina AdaBoost. Allir þessir þrír flokkarar eru tvíkostaflokkarar (e. binary classifiers) og því vel við hæfi fyrir okkar verkefni.

#### 2.3.1 Lógistísk aðhvarfsgreining

Í lógistískri aðhvarfsgreiningu smíðum við líkanið

$$f_{\theta}(x) = \frac{1}{1 + \exp(-\theta^T x)}$$

þar sem við ákvörðum stikana  $\theta$  með því að hámarka logra sennileikafallið

$$\log(L(\theta)) = \sum_{i=1}^n y^{(i)} \log g\left(\theta^T x^{(i)}\right) + \left(1 - y^{(i)}\right) \log \left(1 - g\left(\theta^T x^{(i)}\right)\right). \quad (1)$$

Flokkun gagnanna verður þá

$$\hat{y} = \begin{cases} 1 & \text{ef } g\left(\theta^T \hat{x}\right) \geq 0.5 \\ 0 & \text{annars} \end{cases}$$

Við notum `LogisticRegression` fallið úr `linear model` pakkanum hjá `sci-kit learn`.

### 2.3.2 Stoðvigravél

Við viljum leysa bestunarverkefnið (2) með skorðunum (3)

$$\min_{\theta, \theta_0, \zeta} \frac{\theta^T \theta}{2} + C \sum_{i=1}^n \zeta_i \quad (2)$$

$$\begin{aligned} \text{Skorður: } y^{(i)} \left( \theta^T h(x^{(i)}) + \theta_0 \right) &\geq 1 - \zeta_i, \quad i = 1, \dots, n \\ \zeta_i &\geq 0 \end{aligned} \quad (3)$$

Unnið er með ólínulega stoðvigravél með Gaussískum kjarna(rbf):  $k(u, v) = \exp(-\gamma \|u - v\|_2^2)$ ,  $\gamma > 0$ ,  $\gamma$  er yfirstiki sem við fínstillum ásamt reglunarliðnum  $C$ . Notað var fallið `svm.SVC` úr `sklearn` ásamt því að nota `model_selection.GridSearchCV` fallið úr `sklearn` til að finna gildin á  $C$  og  $\gamma$  sem skiluðu hæstu prófunar- og þjálfunarnákvæmni.

### 2.3.3 Adaboost

Adaboost er sniðug aðferð sem byggist á orðatiltækinu „í krafti fjöldans“. Í flokkunarverkefnum með úttak  $y \in \{-1, 1\}$  er byrjað á því að þjálfra veikan flokkara á gögnunum,  $C_1$ , hann einn og sér stendur sig frekar illa. Næst er skekkjan frá honum reiknuð,  $e_1$ , og hver einasti gagnapunktur sem var metinn rangt fær vogtölu á við  $(1 - e_1)/e_1$ , þannig eru þeir punktar sem voru vitlaust flokkaðir með hærri vigti. Næst er búinn til nýr veikur flokkari  $C_2$  sem flokkar vigtuðu punktana og skekkjan frá nýja flokkaranum reiknar með vogtölunum. Þannig þvingum við nýja flokkarann til þess að einbeita sér betur að þeim punktum sem fyrsti flokkarinn náði ekki að flokka rétt. Áfram er haldið, að setja hærri og hærri vogtölur á þá punkta sem erfitt er að flokka til þess að flokkararnir sem á eftir koma einbeita sér að þeim. Í lokin er skilað flokkaranum

$$C(x) = \text{sign} \left[ \sum_{m=1}^M \log \left( \frac{1 - e_m}{e_m} \right) C_m(x) \right]$$

Þetta er nokkurs konar lýðræðisleg kosning á milli flokkaranna þar sem góðir flokkarar fá meira vægi í sínum atkvæðum [1].

Við notuðum `AdaBoost` flokkarann úr `sci-kit learn` með `DecisionTreeClassifier` með fjögur lauf sem veiku flokkarana. `AdaBoost` hefur tvo yfirstika sem hægt er að stilla af en þeir eru fjöldi veikra flokkara ( $M$  hér að ofan) og lærdómshraði sem minnkar vigtið hvers flokkara með tímanum.

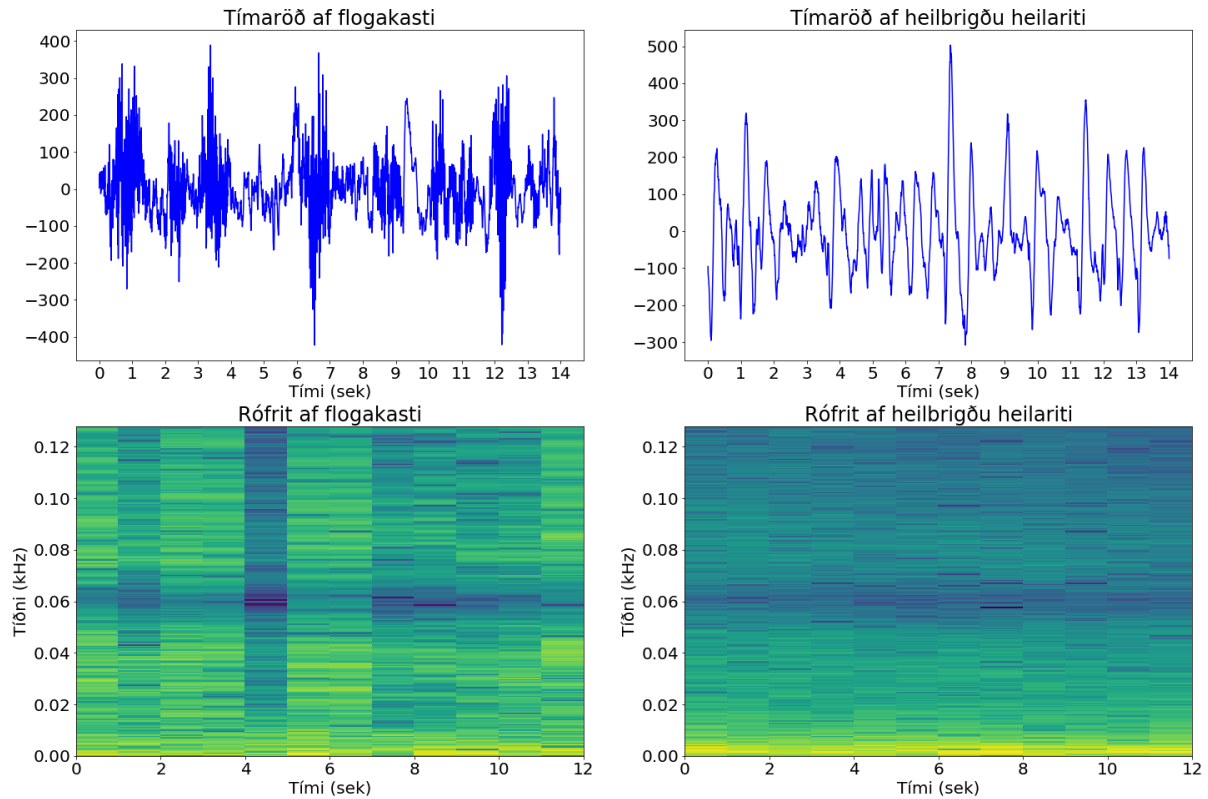
### 2.4 Gæðamat

Þær lýsistærðir sem við notuðum til samanburðar á flokkurunum voru prófunarnákvæmni (e. test set accuracy), sértækni (e. specificity) og næmi (e. sensitivity). Prófunarnákvæmni er grunnmatíð sem oftast er notað, í okkar tilfelli er næmi afar mikilvægt, við viljum komast eins mikið hjá því og við getum að líkanið okkar hleypi ekki flogakasti í gegn án þess að átta sig á því. Í raunverulegum aðstæðum gæti þetta haft lífshættulegar afleiðingar. Sértæknin er ekki eins mikilvæg ein og sér en það er alltaf nauðsynlegt að líta á hana til hliðsjónar við næmnina. Við viljum jú ekki að flokkarinn okkar skapi úlf úlf aðstæður. Lýsistærð Matthew's er oft notuð sem vigti á milli sértækni og næmi og við skoðuðum hana til hliðar. Einnig var fylgst með þjálfunarnákvæmni til þess að reyna að sporna við ofmátun flokkaranna. Það telst almennt ekki gott að ná 100% þjálfunarnákvæmni því þá er flokkarinn frekar farinn að leggja á minnið heldur en að læra. Úr bestu flokkurunum teiknuðum við svo upp ruglingsfylki til að rýna betur í niðurstöður þeirra.

### 2.5 Rófrit og Tauganet

Önnur aðferð til þess að nýta rófs- og tímaeiginleika heilaritanna er með því að skoða rófrít þeirra (e. spectrogram). Aðferðin byggist á Fourier ummyndun tímaraðanna á mismunandi

tímabútum og fyrir eina tímaröð fáum við mynd sem lýsir tíðni raðarinnar á mismunandi tímum [3]. Notast var við `spectrogram` í pakkanum `scipy.signal`. Mynd 2 hér að neðan sýnir tvenn heilarit og samsvarandi rófrit þeirra, eitt með flogakasti og eitt eðlilegt. Þar má sjá hvernig tíðnin er mun stöðugari í eðlilegu heilariti en mjög breytileg í flogakastinu. Sjá má á myndum heilbrigða heilaritsins hvernig rófritið geymir ekki upplýsingar um sveifluvídd, heldur aðeins um tíðni og tíma.



Mynd 2: Tímaraðir og Rófrit

Með því að umbreyta sérhverri tímaröð í rófrit þá getum við litið svo á að hvert af okkar 540 tilvikum hefur eina 23 laga mynd þar sem hvert lag er rófrit úr einni rás. Þessar myndir getum við notað til að þjálfra földunartauganet. Földunarnet hafa hvert sinn strúktúr, þ.e. hversu mörg hulin lög það hefur, hversu marga hnúta, hversu stóran földunarglugga o.s.frv. Við prófuðum ýmis mismunandi strúktúra sem virkuðu misvel. Við bárum saman strúktúrana með tilliti til prófunarnákvæmni, prófunartapi, næmi og sértækni. Í þjálfuninni fylgjumst við með prófunartapinu og ef það batnar ekki í þrjú tímabil (e. epochs) í röð þá hættum við þjálfuninni og sættum okkur við núverandi líkan. Gagnasettið hér er ekki voða stórt, hvert inntak er fremur stórt (cirka  $256 \times 14 \times 23$  víð mynd) en við höfum aðeins 540 inntök. Tauganet virka best ef fyrirbyggur stórt gagnasett, því er líklegt að við munum ofmáta tauganetið.

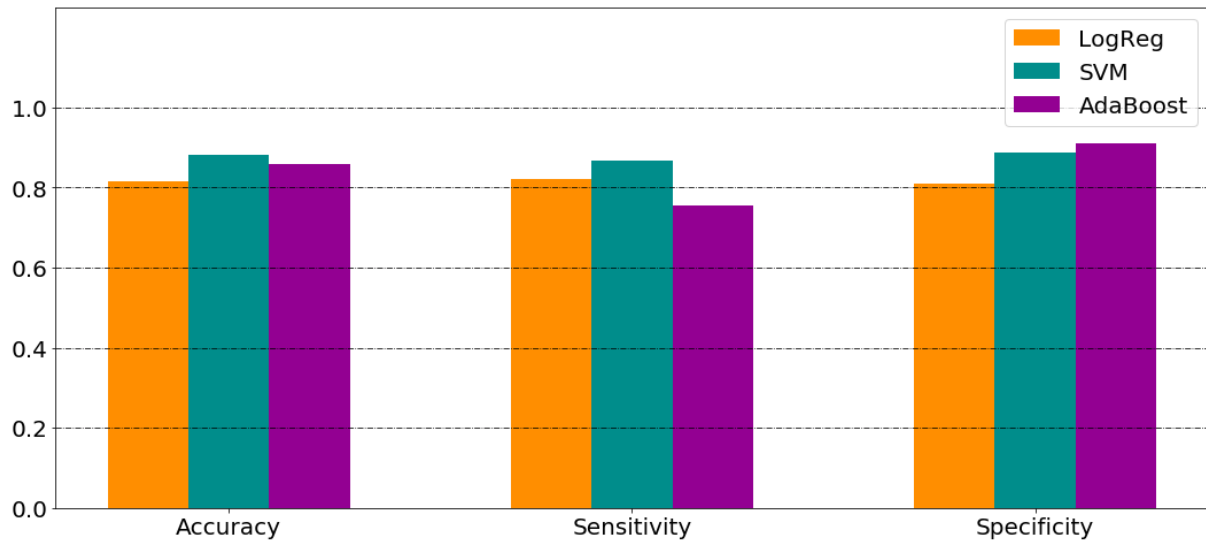
## 2.6 Einstaklingsmiðuð greining

Eftir að hafa borið saman flokkarana á öllu gagnasafninu þá völdum við þann besta og athuguðum hvernig hann stóð sig í því að greina einn einstakling í einu. Gögnunum var þá skipt niður eftir einstaklingum, flokkarinn þjálfaður á öllum einstaklingum nema einum og svo prófaður á öllum tilvikum þess einstaklings. Nákvæmni, sértækni og næmi flokkarans voru reiknuð og svo allt endurtekið fyrir hvern einstakling. Einstaklingarnir hafa mismargar mælingar og mismörg tilvik af flogaköstum.

### 3 Niðurstöður

#### 3.1 Grindarleit

Framkvæmd er grindarleit (e. grid search) á þremur flokkurum SVM, Adaboost og Logistic regression í þeim tilgangi að finna þá yfirstika sem skila bestu prófunarnákvæmni ásamt næmi og sértækni.

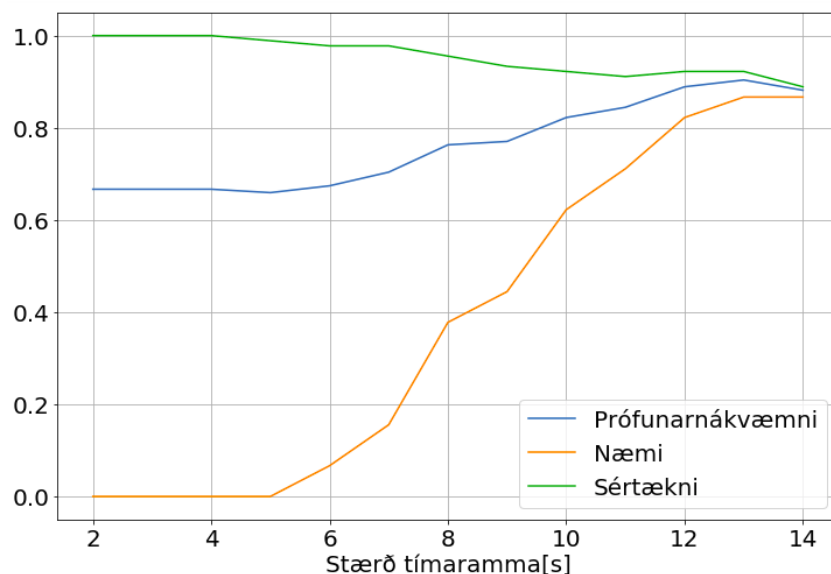


Mynd 3: Samanburður á prófunarnákvæmni, næmi og sértækni fyrir flokkarana með sínum bestu yfirstikum fyrir stærð tímarammans sem 14 sekúndur, Fáum eftirfarandi stika **LogReg**:  $C = 20$ ; **SVM**:  $C = 8$ ,  $\gamma = 0.0015$ ; **AdaBoost**: `learning_rate=0.5`, `n_estimators=150`

Við sjáum að stoðvigravél stendur sig best á þeim vígvöllum sem við höfum helstan áhuga á, prófunarnákvæmni og næmi. Við ákveðum því að stoðvigravél með þá yfirstika sem ákvarðaðir voru með grindarleitinni ( $C = 8$  og  $\gamma = 0.0015$ ) sé besti flokkarinn fyrir okkar verkefni. Við skoðuðum næst hvernig mismunandi stærð á tímarömmum hefur áhrif á flokkarann.

#### 3.2 Tímarammi

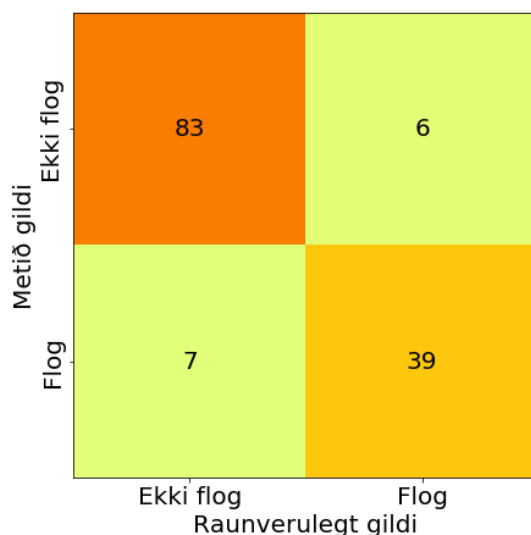
Við skoðuðum alla mögulega tímaramma og mátuðum stoðvigravélina okkar við samsvarandi gögn og reiknuðum lýsistærðirnar fyrir hvern tímaramma. Niðurstöðurnar má sjá hér að neðan í mynd 4.



Mynd 4: Prófunarnákvæmni, næmi og sértækni stoðvigravéla sem fall af stærð tímaramma. Notað er stoðvigravél með yfirstika  $C = 8$  og  $\gamma = 0.0015$

Við höfum mestan áhuga á prófunarnákvæmni og næmi og sjáum að við fáum hágildi á þeim stærðum þegar stærð tímarammans er í 13 sekúndum. Þegar stærð tímarammans er 13 sekúndur fáum við 90.37% í prófunarnákvæmni, 86.67% í næmi og 92.22% í sértækni sem er ágætlega gott.

Með réttu yfirstika og rétta stærð af tímaramma þá höfum við sætt okkur við lokaflokkarann okkar og mynd 5 hér að neðan sýnir ruglingsfylki hans.



Mynd 5: Ruglingsfylki Stoðvigravélarinnar

Þessi niðurstaða er vel ásættanleg, við athugum nú hvernig þessi flokkari hagar sér þegar hann greinir einn einstakling í einu.

### 3.3 Einstaklingsmiðuð greining

Hér á mynd 6 má sjá hvernig útvaldi flokkarinn okkar stóð sig í því að greina heilarit hvers einstaklings fyrir sig.

Patient	# Seizures	# non Seizures	Accuracy	Sensitivity	Specificity	Matthew's
1	7	11	0.833	0.571	1.000	0.670
2	2	15	1.000	1.000	1.000	1.000
3	7	5	0.917	1.000	0.800	0.837
4	4	14	0.778	0.250	0.929	0.236
5	4	15	0.947	1.000	0.933	0.864
6	10	3	0.846	0.800	1.000	0.693
7	3	4	1.000	1.000	1.000	1.000
8	5	5	0.800	0.800	0.800	0.600
9	4	4	0.625	0.250	1.000	0.378
10	7	9	0.812	0.857	0.778	0.630
11	3	13	0.812	1.000	0.769	0.620
12	27	4	0.290	0.222	0.750	-0.022
13	10	1	0.364	0.300	1.000	0.194
14	8	6	0.571	0.250	1.000	0.354
15	19	3	0.273	0.211	0.667	-0.101
16	1	2	1.000	1.000	1.000	1.000
17	3	6	0.889	0.667	1.000	0.756
18	6	11	0.588	0.333	0.727	0.064
19	3	10	0.769	0.667	0.800	0.426
20	7	8	0.533	0.000	1.000	0.000
21	4	9	0.692	0.000	1.000	0.000
22	3	8	0.909	1.000	0.875	0.810
23	7	1	0.875	1.000	0.000	0.000
24	16	3	0.895	0.875	1.000	0.725
MEAN			0.721	0.602	0.833	0.469

Mynd 6: Einstaklingsmiðuð greining stoðvigravélarinnar

Niðurstöðurnar eru ekki eins góðar og við vonuðumst eftir. Við náum aðeins 72.1% meðalprófunarnákvæmni og aðeins 60.2% meðalnæmi. Þetta telst alls ekki nógu gott til þess að nýta flokkarann fyrir alvöru. Með rýnun í töfluna að ofan má sjá að oft vill flokkarinn bara giska á annan flokkinn alltaf. Ástæðuna fyrir þessari lélegu frammistöðu má líklega rekja til skölunar á gögnunum. Nú eru stoðvigravélar mjög næmar fyrir skölun og gögnin hafa mikla dreifni svo skölunin getur verið mismunandi eftir því hvert þjálfunarsafnið er og þessir fáu prófunarpunktar gætu skalast furðulega. Við prófuðum bæði að sleppa sköluninni og breyta yfir í MinMax skölun en það skilaði aðeins verri niðurstöðum. Einnig voru hinir flokkararnir, lógistíska aðhvarfsgreiningin og AdaBoost með bestu yfirstikunum úr grindarleitinni prófaðir en þeir skiluðu einnig verri niðurstöðum.

### 3.4 Tauganet

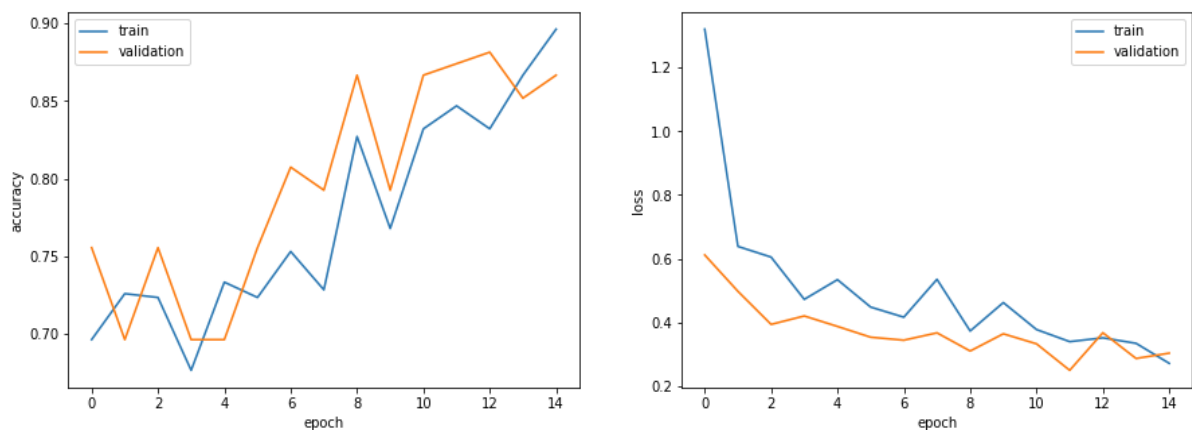
Eftir margar prófanir á mismunandi strúktúrum földunarnetsins þá enduðum við að vinna með eftirfarandi strúktúr sem má sjá á mynd 7.

Í orðum þá byrjuðum við á tveimur 60 hnúta földunarlögum með földunarglugga að stærð  $6 \times 2$ , ástæðan fyrir þessari afbrigðilegu gluggastærð er að upphaflegu myndirnar okkar eru  $257 \times 13$  að stærð svo við viljum safna upplýsingum um fleiri tíðnibil á þrengri tímabilum. Eftir fyrstu földunarlögin kemur safnlag með  $2 \times 2$  glugga, tvö 120 hnúta földunarlög með  $3 \times 3$  glugga, safnlag, tvö 180 hnúta földunarlög, safnlag og svo þrjú full tengd lög með 1000, 200 og 100 hnútum með úrfellingum á milli.

Þetta tauganet náði prófunarnákvæmni upp á 86.7% og tap upp á 0.3033 eftir að hafa lært í 11 tímabil (e. epochs) en yfirlit um þjálfunarsögu netsins má sjá á mynd 8.

Layer (type)	Output Shape	Param #
conv2d_52 (Conv2D)	(None, 257, 13, 60)	16620
conv2d_53 (Conv2D)	(None, 257, 13, 60)	43260
max_pooling2d_24 (MaxPooling)	(None, 85, 6, 60)	0
conv2d_54 (Conv2D)	(None, 85, 6, 120)	64920
conv2d_55 (Conv2D)	(None, 85, 6, 120)	129720
max_pooling2d_25 (MaxPooling)	(None, 42, 3, 120)	0
conv2d_56 (Conv2D)	(None, 42, 3, 180)	194580
conv2d_57 (Conv2D)	(None, 42, 3, 180)	291780
max_pooling2d_26 (MaxPooling)	(None, 21, 1, 180)	0
flatten_7 (Flatten)	(None, 3780)	0
dense_28 (Dense)	(None, 1000)	3781000
dropout_14 (Dropout)	(None, 1000)	0
dense_29 (Dense)	(None, 200)	200200
dropout_15 (Dropout)	(None, 200)	0
dense_30 (Dense)	(None, 100)	20100
dense_31 (Dense)	(None, 1)	101
Total params: 4,742,281		
Trainable params: 4,742,281		
Non-trainable params: 0		

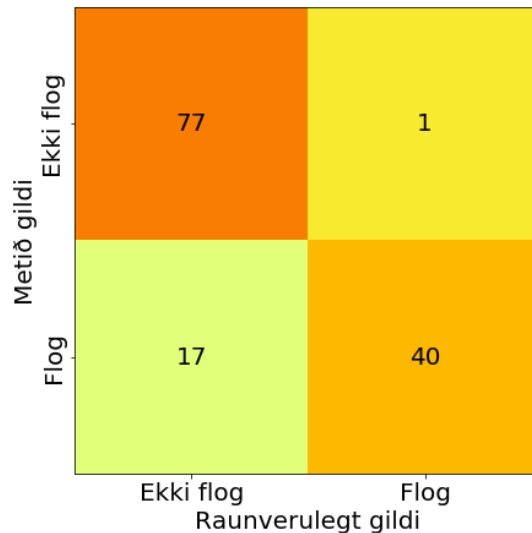
Mynd 7: Strúktúr Tauganetsins



Mynd 8: Lærdómur Tauganetsins, nákvæmni til vinstri og tap til hægri

Nákvæmnin flóktar heldur mikið sem er eðlilegt með svo lítið gagnasett en tapið hegðar sér fremur vel. Við sjáum svo ruglingsfylki tauganetsins hér að neðan á mynd 9.





Mynd 9: Ruglingsfylki Tauganetsins

Eins og nefnt er hér að ofan þá er næmi mikilvægasta lýsistærðin okkar og tauganetið náði næmi upp á 97.56%, sem er frábært. Við sjáum af ruglingsfylkinu hér að ofan að það er aðeins eitt tilvik þar sem tauganetið missir af flogakasti. Þrátt fyrir þessa háu næmi þá borgum við ekki stórt gjald í sértækninni á móti því hún er upp á 81.91%, Matthew stuðullinn er upp á 0.74 sem er ágætt sérstaklega fyrir svona háa næmi.

Við prófuðum að nota tauganetið okkar sem einstaklingsmiðaðann flokkara en það fór ekki vel, meðal nákvæmnin endaði í 60.6%, næmnin í 45.7% og sértæknin í 65.2%. Svo virtist sem tauganetið hallaðist að því að giska alltaf á ekki flog.

## 4 Samantekt

Okkur hefur tekist að smíða tvo ólíka flokkara, stoðvigravél og földunarnet, sem nema flogaköst í heilaritum með annars vegar 90.37% prófunarnákvæmni og 86.67% næmi með stoðvigravélinni og hins vegar 86.7% prófunarnákvæmni og 97.56% næmi með tauganetinu. Flokkararnir tveir notuðu ólík auðkenni sem inntaksgögn sem þó bæði byggðu á tíðni- og tímaupplýsingum. Hægt er að vinna önnur auðkenni úr tímaröðum og mögulegt er að þau skili betri árangri í flokkun en gert var hér. Bæði stoðvigravélin og tauganetið stóðu sig leiðinlega illa í því að greina einstaklingsbundin flogaköst, enda er það erfitt verk því flogaköst eiga það til að koma ólíkt fram hjá ólíku fólki. Það er því eitthvað sem þyrfti að leggja betri áherslu á í áframhaldandi vinnslu á þessu verkefni.

## Heimildir

- [1] J. FRIEDMAN, T. HASTIE, AND R. TIBSHIRANI, *The elements of statistical learning*, vol. 1, Springer series in statistics New York, 2001.
- [2] J. MALMIVUO AND R. PLONSEY, *Bioelectromagnetism. 13. Electroencephalography*, 01 1995, pp. 247–264.
- [3] J. NUNEZ-IGLESIAS, S. VAN DER WALT, S. WALT, AND H. DASHNOW, *Elegant SciPy: The Art of Scientific Python*, O'Reilly Media, Inc., 2017.
- [4] Y.-R. L. SEUNG-HYEON OH AND H.-N. KIM, *A novel eeg feature extraction method using hjorth parameter*, International Journal of Electronics and Electrical Engineering, (2014).

- [5] A. H. SHOEB, *Application of machine learning to epileptic seizure onset detection and treatment*, Massachusetts Institute of Technology, (2009).