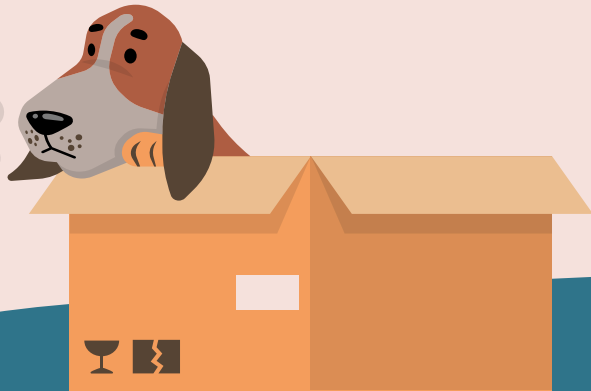


Pawpularity

Demi Soetens, Bente van Katwijk, Ilse Feenstra, Cass Maes





‘Pawpularity’ van zwerfdieren



Wereldwijd miljoenen zwerfdieren op straat of in asiel



Foto's voor asiel Maleisië

- Welke doen het het beste?



Pawpularity score: hoe vaak een pagina bekeken wordt

- Tussen 0 en 100



Doel: het creëren van een model dat de Pawpularity score van foto's voorspelt



3

A close-up photograph of a small, dark-furred puppy with light-colored markings on its chest and paws, looking directly at the camera. The puppy has large, dark eyes and floppy ears. It is being held by a person wearing a blue shirt, whose arm is visible on the right side of the frame. The background is a blurred green, suggesting an outdoor setting.

- ## Binaire data

[illegible]

Pawpularity score: 100

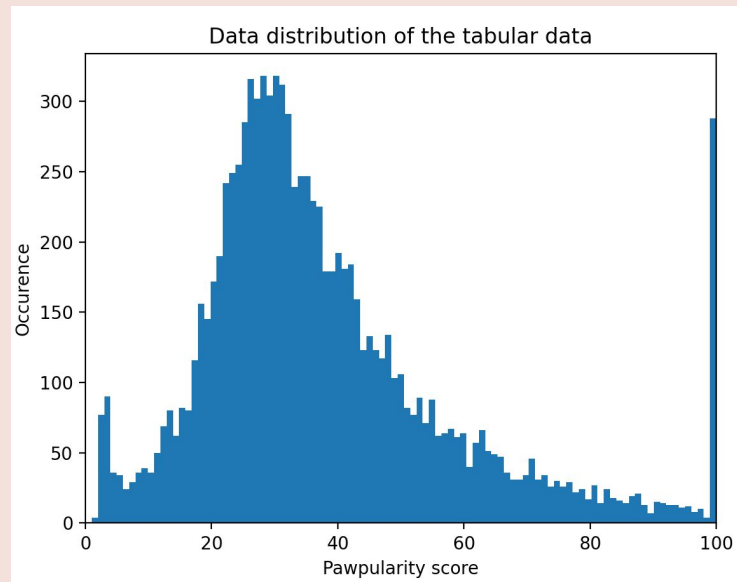


Pawpularity score: 14



Preprocessing van de data

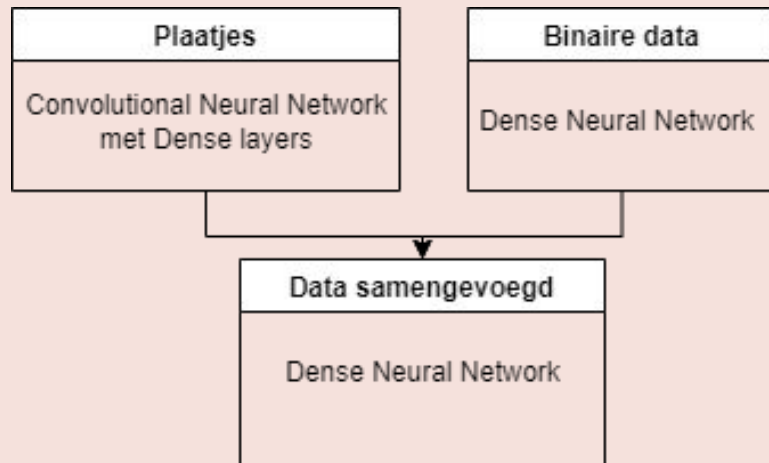
- Images: 64 x 64 x 3
- Image normalization
- Outliers
- Data augmentation



Distributie van de Pawpularity score van de trainings data

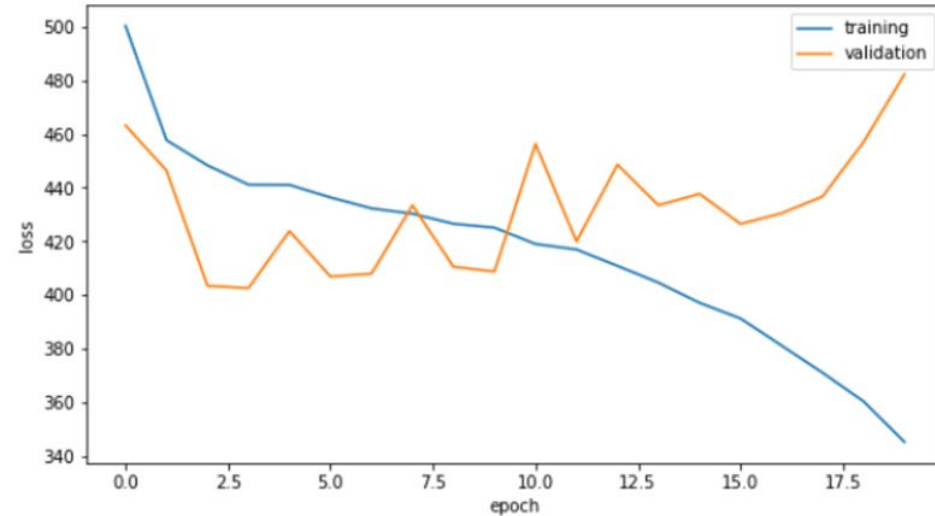
Basis model

- Plaatjes: Convolutional neural network met Dense layers
- Binaire data: Dense neural network
- Loss: mean squared error (MSE)
- Metric: root mean squared error (RMSE)

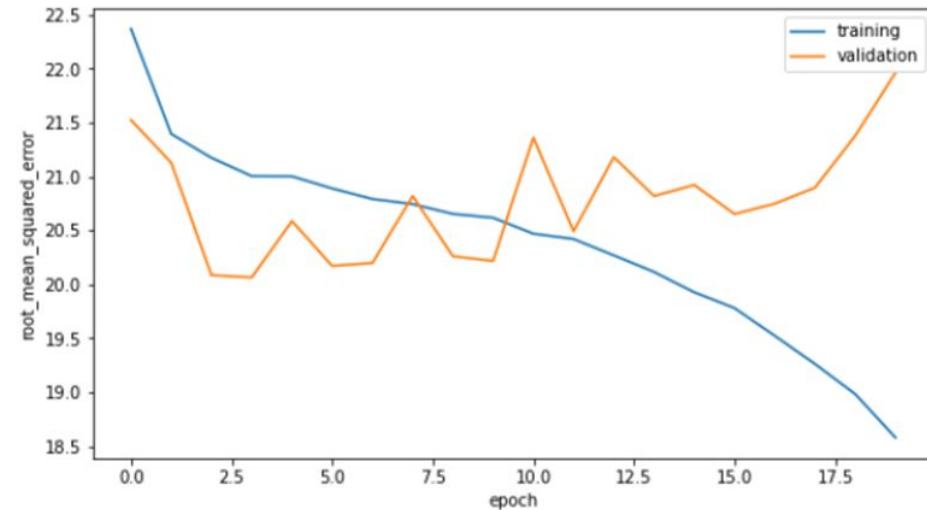


Resultaat basis model

Model loss of basic model



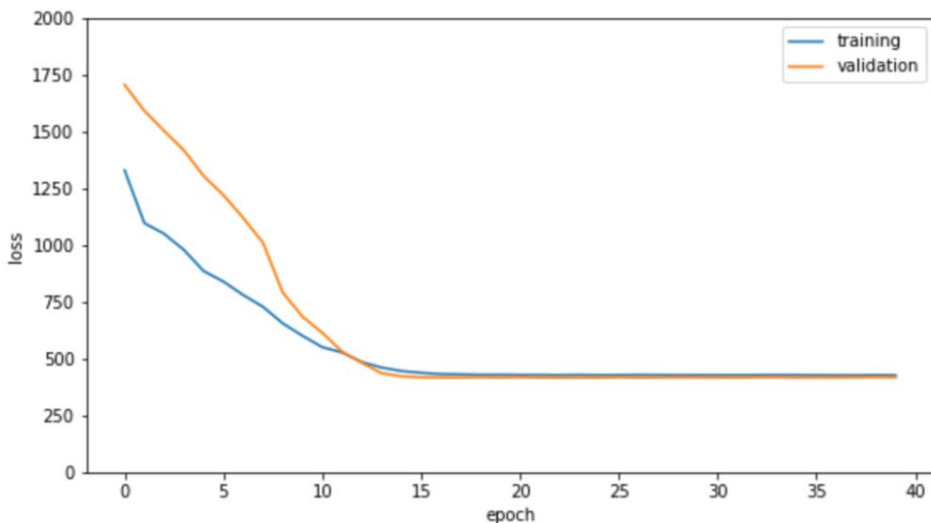
A. Mean squared error



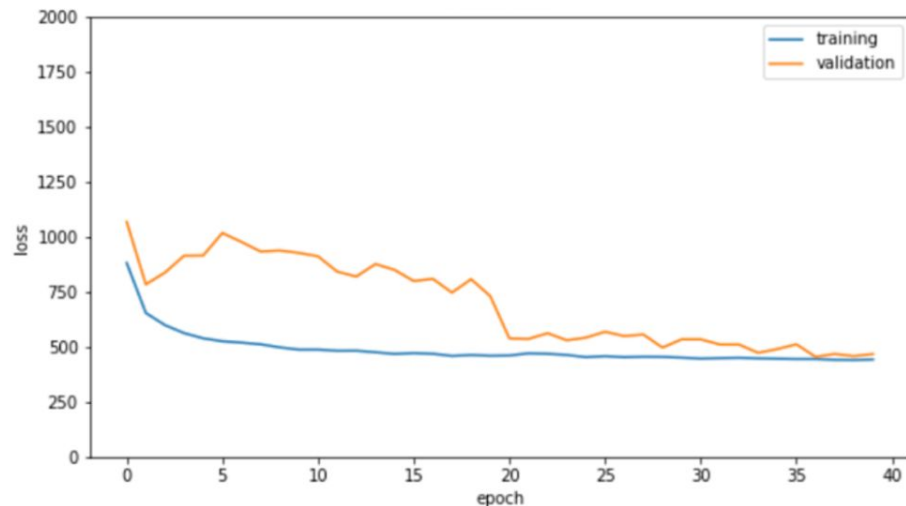
B. Root mean squared error

Resultaten dropout

Comparison of the model loss of high dropout rate vs. a reduced dropout rate



A. Model with high dropout rate

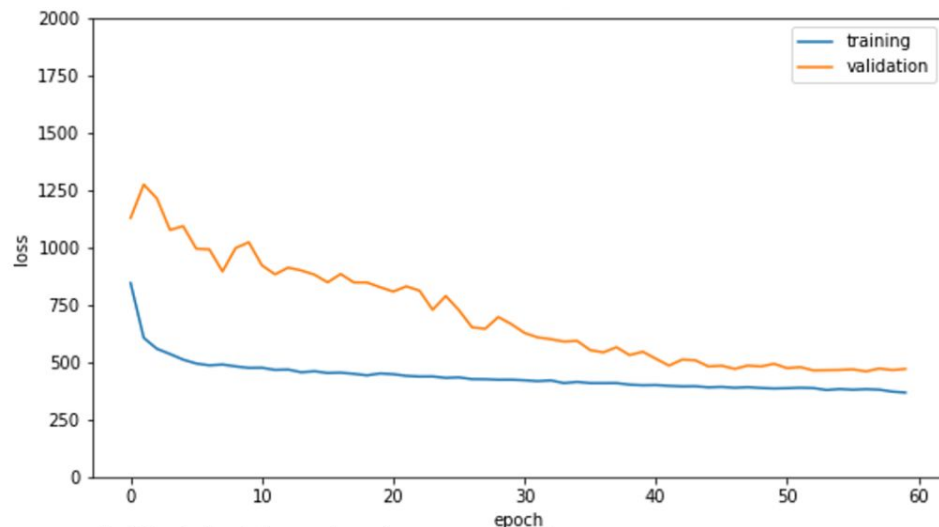


B. Model with reduced dropout rate

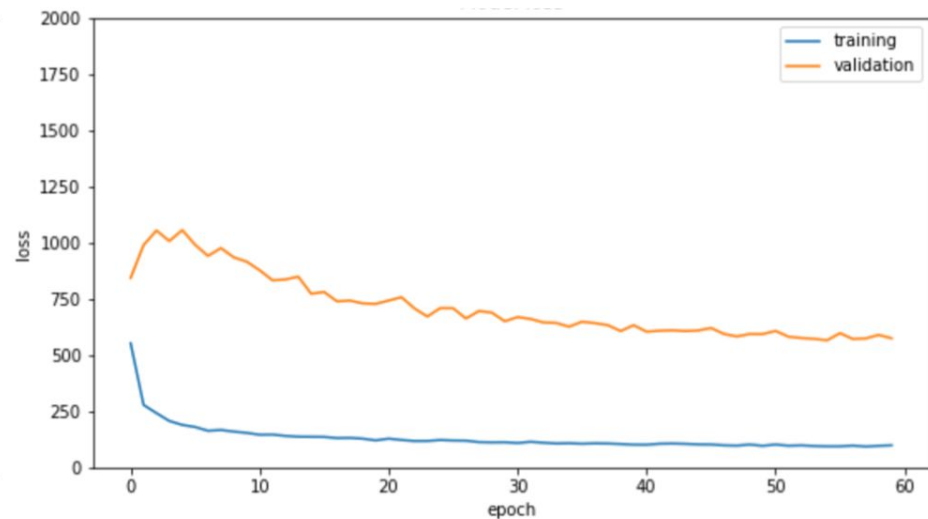
Leercurve van model (MSE) zonder (links) vs met (rechts) dropout

Batch normalization

Comparison of the model loss with vs without batch normalization



A. Model without batch normalization

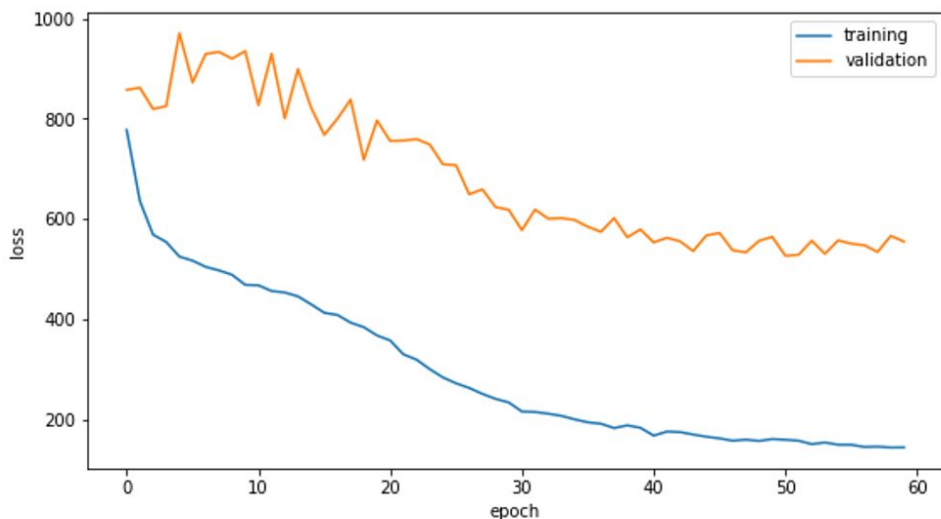


B. Model with batch normalization

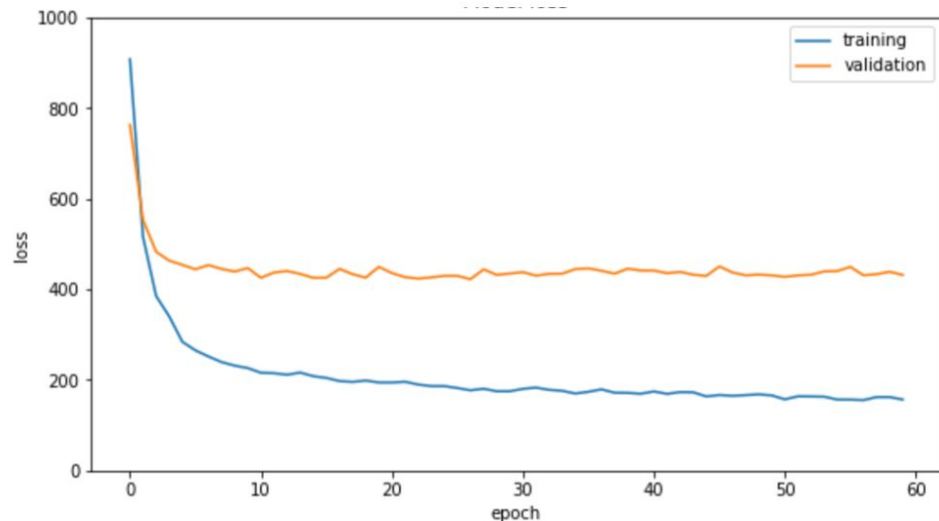
Leercurve van model (MSE) zonder (links) vs met (rechts) batch normalization

Regularization term

Comparison of the model loss with vs. without regularization term



A. Model without regularization term



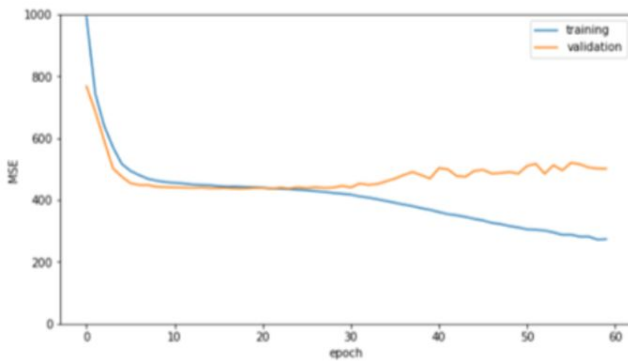
B. Model with L2 regularization term

Leercurve van model (MSE) zonder (links) vs met (rechts) regularization term

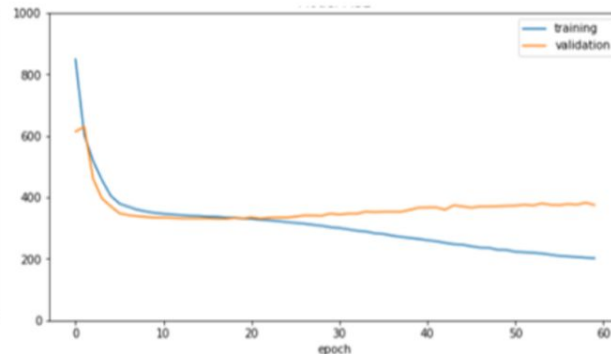
Outliers & data augmentation

- Outliers bij Pawpularity score 100
- Data augmentation:
 - horizontal flip, shear range 20% en rotation range 90°

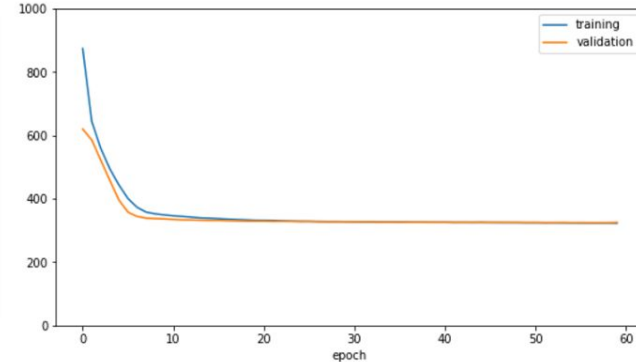
Comparison of the model loss (MSE) with vs without removal of outliers and data augmentation



A. Previous model



B. Removal outliers

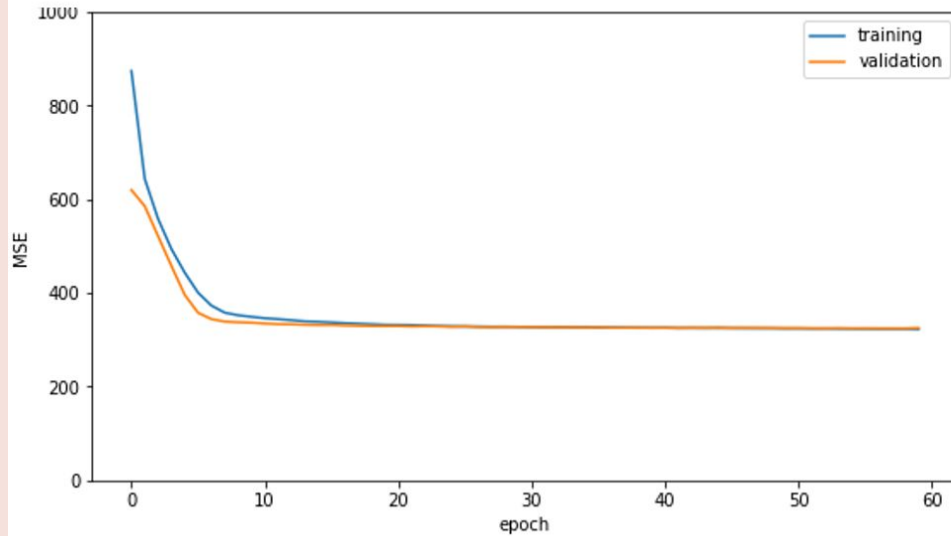


C. Removal outliers and data augmentation

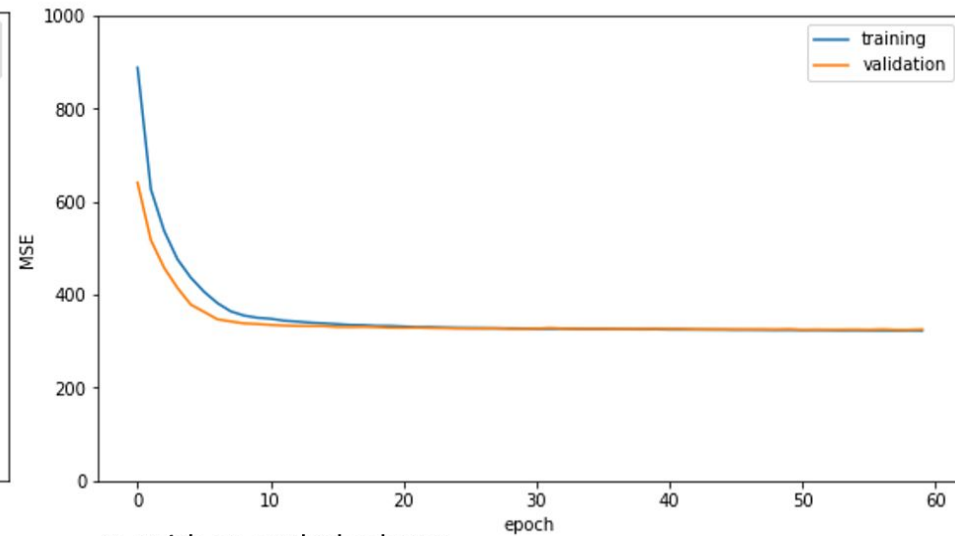
Leercurve van model (MSE) met outliers (links) vs zonder outliers (midden) vs zonder outliers en data augmentation (rechts)

Verwijderen lagen

Comparison of model loss with full tabular network vs simple tabular network



A. With full tabular network

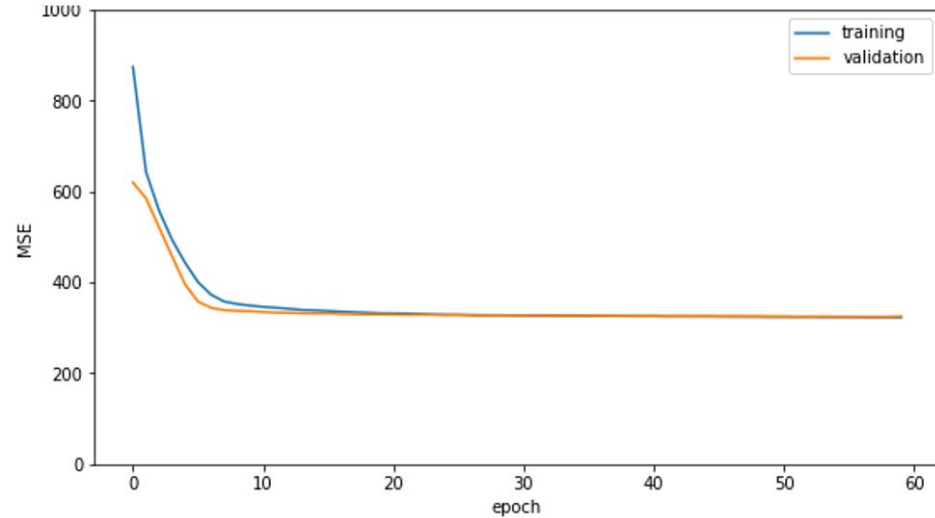


B. With one tabular layer

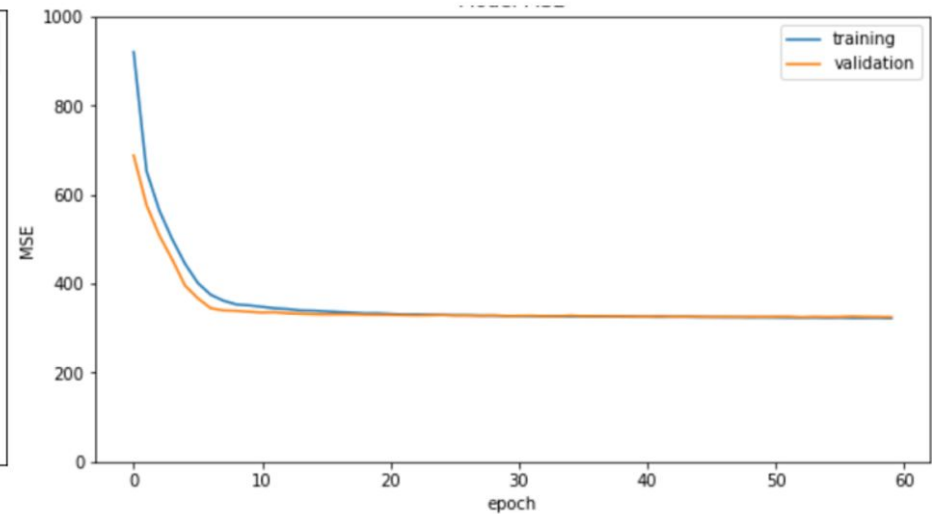
Leercurve van model (MSE) met hele netwerk (links) vs netwerk met lagen verwijderd

Extra convolutional lagen

Comparison of model loss with and without one extra layer per convolutional layer



A. Without extra convolutional layers

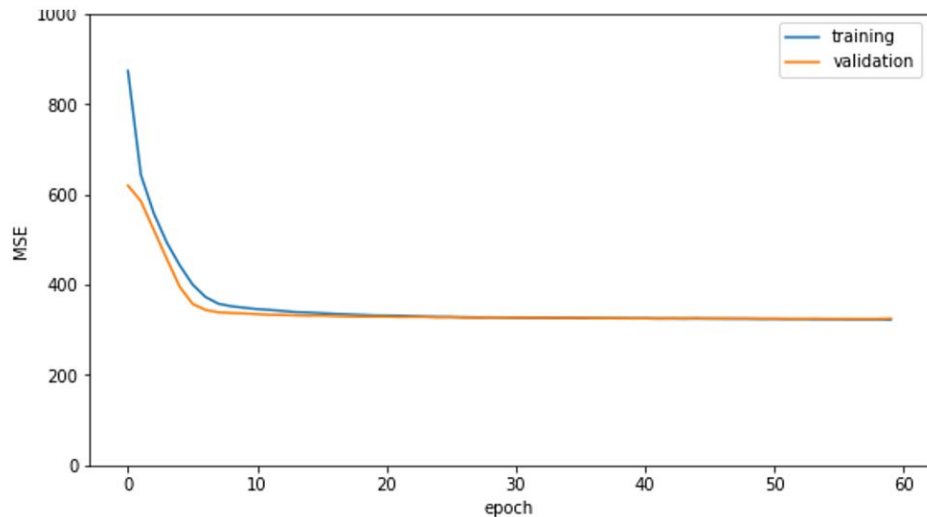


B. With one extra convolutional layers

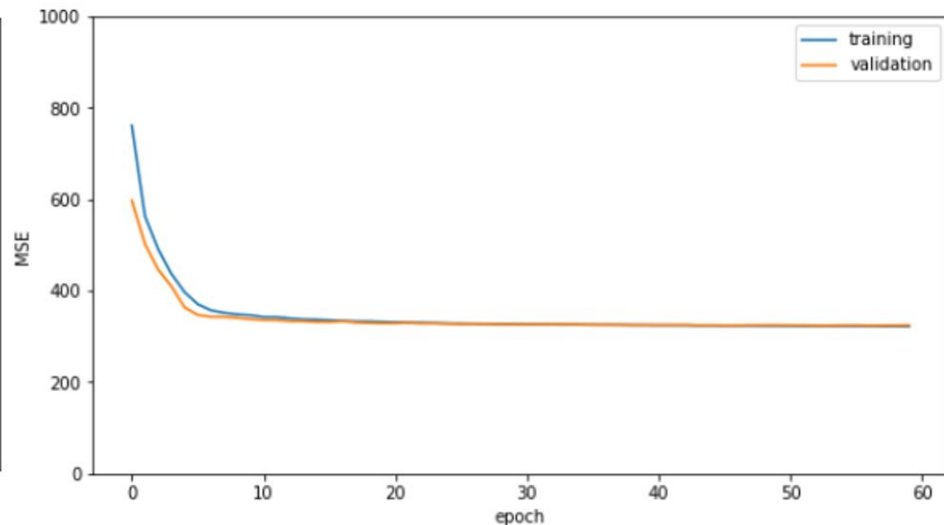
Leercurve van model (MSE) met meerdere extra lagen (links) vs netwerk met een extra laag

Extra hidden nodes

Comparison of model loss with and without extra hidden nodes



A. Without extra hidden nodes



B. With extra hidden nodes

Overzicht uiteindelijke model



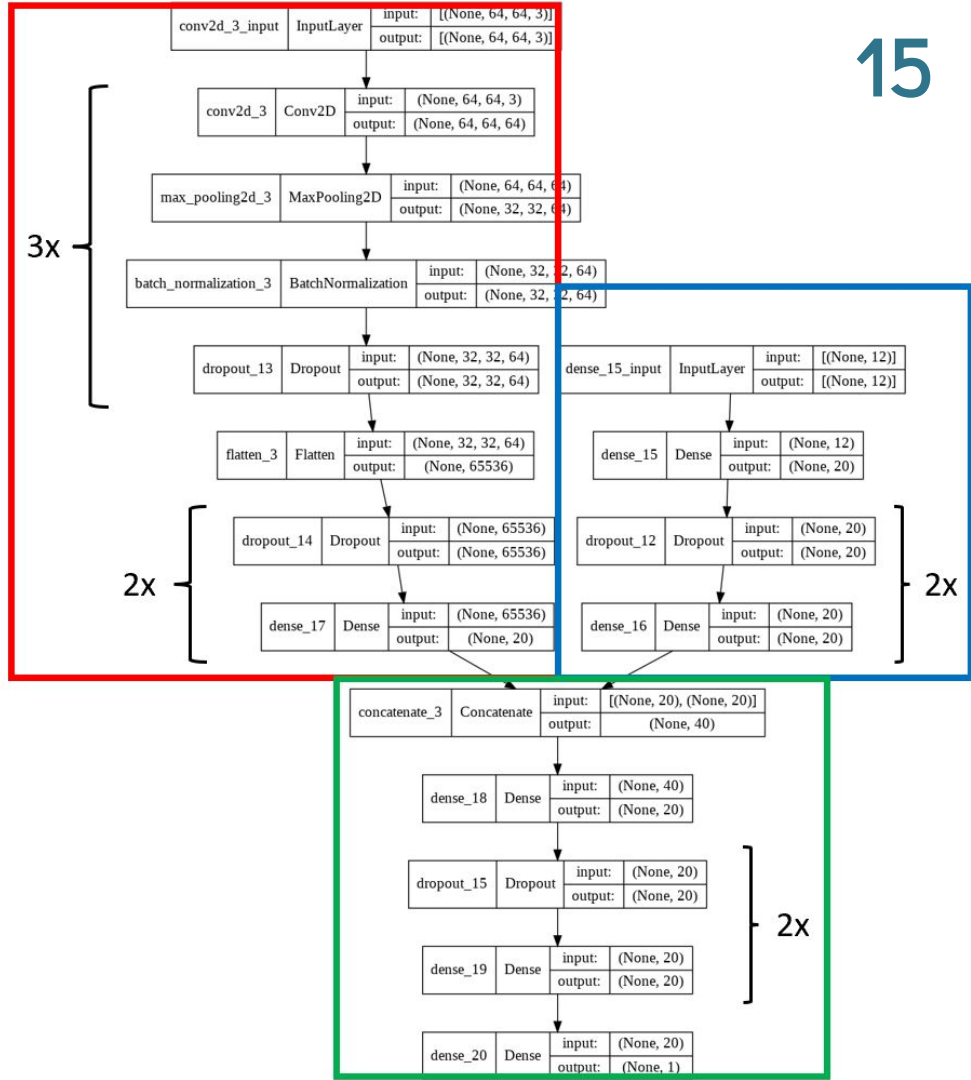
Tabular network



Convolutional network

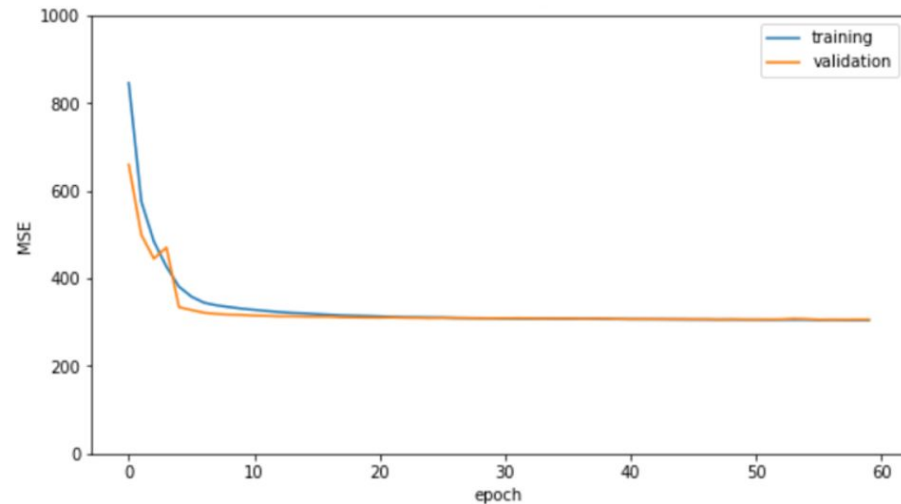


Concatenated network

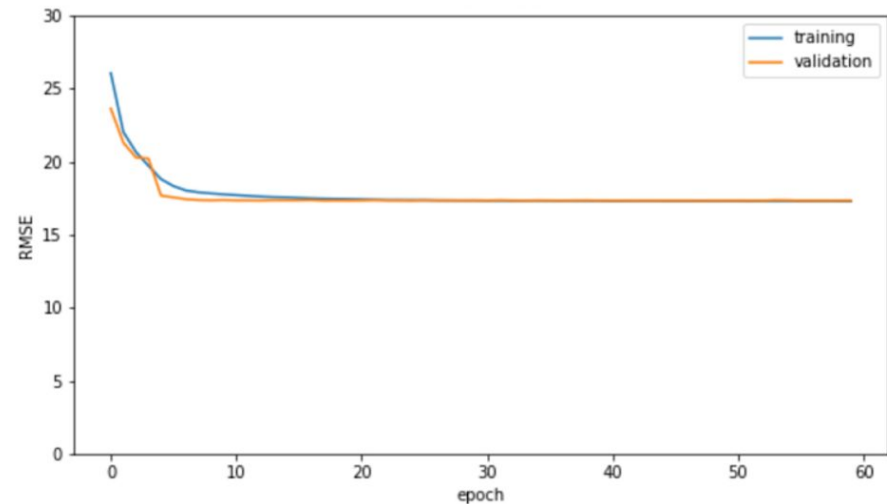


Prestatie uiteindelijke model

Model loss of final model







A. MSE






B. RMSE

Model pipeline

Do's

-  Dropout
-  Batch normalization
-  Regularization terms
-  Lineaire output 0 - 100

Don'ts

-  Verwijderen lagen in tabular netwerk
-  Extra convolutional lagen toevoegen
-  Meer hidden nodes toevoegen

Conclusie




















RMSE validation: 17.24



Plek 1 in kaggle (public leaderboard)

■ In the money
 ■ Gold
 ■ Silver
 ■ Bronze

#	Team Name	Notebook	Team Members	Score ?	Entries	Last
1	🏍️ & 🧊 & 🍄 & 🦋 & 🐸		    	17.48600	191	17d
2	newbees		   	17.59291	187	19d
3	[RAPIDS.AI] Takahagi [Rist]		  	17.59617	96	17d
4	NaN		   	17.61170	73	17d
5	hyd			17.61640	89	17d



Hoe nu verder?

- 🐾 Optimizer
- 🐾 Transfer network
- 🐾 Ensemble of models
- 🐾 Eigen features bouwen
- 🐾 Log transformatie
- 🐾 Parameters automatisch runnen





Bedankt voor jullie aandacht!

Vragen?



