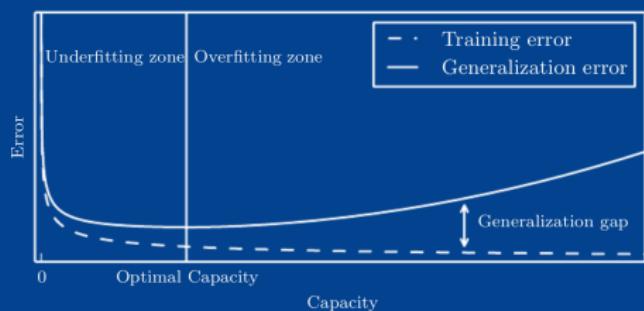




LESSON 7: Model-capacity, Under/over-fitting, Generalization

CARSTEN EIE FRIGAARD

SPRING 2021



"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E ." — Mitchell (1997).

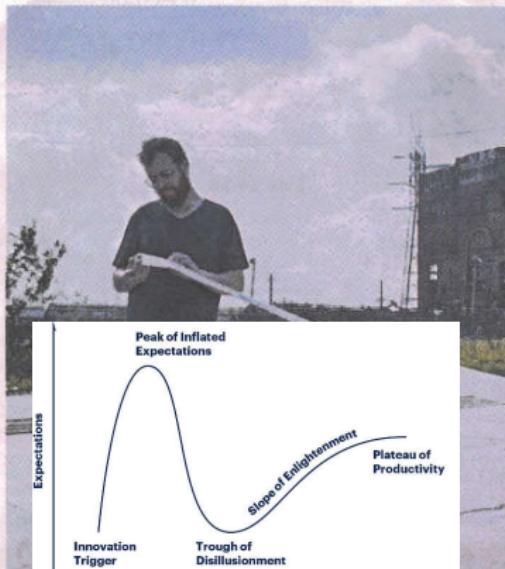
L07: Model-capacity, Under/over-fitting, Generalization

Agenda

- ▶ Spørge-minutter..
- ▶ Demo af ML-systemer og klasse diskussion.
- ▶ Model Capacity,
- ▶ Under/over-fitting,
Exercise: [L07/capacity_under_overfitting.ipynb](#)
- ▶ Generalization Error,
Exercise: [L07/generalization_error.ipynb](#)







Werketlavisen

Bøger

Anna Rossinelli Interview med den britiske romanforfatter Ross Goodwin om sin debutroman *Bøger*, der følger en insekthjerne på en roadtrip.

Kathrine Henningsen Interview med den danske forfatter Kathrine Henningsen om sin debutroman *Bøger*.

Ross Goodwin Illustration af Ross Goodwin.

Bøger

Interview. Den unge romanforfatter Ross Goodwin har fået ført en kærlig insekthjerne, som har skrevet en surrealistisk historie.

På roadtrip med en insekthjerne

Det er ikke en almindelig roadtrip, der er beskrivet i den britiske debutroman *Bøger* af Ross Goodwin. Det er en rejse, der foregår i hjernen til en insekthjerne, der skal tilbage til sit hjem i det sydlige England. I baggrunden ses en del af en by i New Orleans, hvor den britiske forfatter bodde i et hus sammen med sin kone og deres tre børn.

Det er ikke en almindelig roadtrip, der er beskrivet i den britiske debutroman *Bøger* af Ross Goodwin. Det er en rejse, der foregår i hjernen til en insekthjerne, der skal tilbage til sit hjem i det sydlige England. I baggrunden ses en del af en by i New Orleans, hvor den britiske forfatter bodde i et hus sammen med sin kone og deres tre børn.

Det er ikke en almindelig roadtrip, der er beskrivet i den britiske debutroman *Bøger* af Ross Goodwin. Det er en rejse, der foregår i hjernen til en insekthjerne, der skal tilbage til sit hjem i det sydlige England. I baggrunden ses en del af en by i New Orleans, hvor den britiske forfatter bodde i et hus sammen med sin kone og deres tre børn.



På roadtrip med en insekthjerne

Af MIKKEL BORIS

T 2016 slog computeren AlphaGo den 18-dobbelte verdensmester i brætspillet Go, Lee Sedol. Go er et kompliceret og abstrakt spil, som kræver intuition og kreativitet, men det kunstige intelligens vandt med en række innovative træk overlegen. Undervis i støller roede kommentatorerne,

Goodwin til Weekendavisen fra sin lejlighed i Los Angeles.

Inden køreturen havde han brugt måneder på at træne maskinen. Han satte den til at læse et stort korpus af moderne litteratur fra hele verden, så den kunne lære at skrive af de store forfattere.

»Det fungerer ligesom autokorrekturen på din telefon, bare klogere og trænet på en mere litterær kilde. Den skriver bogstav for bogstav, så den har lært sig selv at formulere det næste,

når du dekonstruerer dem. Efter at have læst den i ét stræk og fået turen lidt på afstand har romanen fået en universalitet, så jeg kan projicere mine egne oplevelser ind i teksten,« uddyber Goodwin.

– *Du har beskrevet projektet som at leve en insekthjerne at skrive. Hvad betyder det?*

»Jeg forsøgte at pointere, at maskinen ikke er på niveau med den menneskelige hjerne. Et artificielt netværk net er en algoritme, der er lavet, som vi troer hjernen fungerer. Jeg synes,

»Det er et forsøg på at skabe en ny brugerflade for at skrive. På en måde har jeg jo skrevet en roman med en bil.« fortæller Ross Goodwin om sin AI-forfattede bog *1 the road*

Hvad siger pressen?

BLOG | Erik David John

GPT-3 er ikke stærk AI

Erik David Johnson Tirsdag, 9. marts 2021 - 15:00

11



Sidste år så vi GPT-3 melde sin ankomst på AI-scenen. For de af man med GPT-3 producere meningsfulde tekster i natursprog (n) uset niveau. Autogenererede sætninger og mindre artikler var som menneskelig forfatter, og endnu en gang blev AI fremhævet som vej til at ændre verden.

Ikke så imponerende som det lyder

Sagen er bare at det ikke er så revolutionerende som skaberne af GPT-3 et øjeblik ville man forstå at der



Om Erik David John

Erik er Principal AI Strateg. Han arbejder med forretning

Robotter kan nu lugte og mærke (næsten) ligesom mennesker

Sanserne er nødvendige, når robotter skal begå sig i menneskets kaotiske verden.

AF EBSEN HARDEMBERG 26. JUNI 2019 - BEMÆRK: ARTIKLEN ER MERE END 30 DAGE GAMMEL

[?] FORKLAR ORD [•] STØRRE TEKST [o] LÆS OP

I Odense er firmaet AmiNIC ved at udvikle en ny robot, med en ganske særlig evne: lugtesans. Robotten skal bruges til at lugte til fisk og afgøre, om fisken er frisk eller ej. Det kan hjælpe restauranter og fædrevareindustrien til kun at smide de råvarer ud, som er fordærvede.

På den anden side af Atlanteren, i Cambridge, på MIT-universitets afdeling for IT og



ING/VERSION2
NYHEDER BLOGS DEBAT JOB SEKTIONER ▾ MERE ▾ IT-TALENT INFOSECURITY TIP OS ANONYMT

Banker vil vurdere boligpriser med AI



(Illustration: Puttachit Kumkrom / Bigstock)

- Mest debatterede Mest læste
- 1984, Ghostbusters og ...
 - Regningen for Smittestop-appen stiger
 - Frankrig og Holland stætter EU i kamp mod amerikanske IT-gigantter
 - Forslag om eget IT-kontrol i det offentlige: 'Det her stykke arbejde SKAL løves'

WEBINARS AND WHITEPAPERS

- WEBINAR Objektivitet, flexibilitet og muligheder for bazierte løsninger - og



A computer vision system to monitor the infestation level of Varroa destructor in a honeybee colony

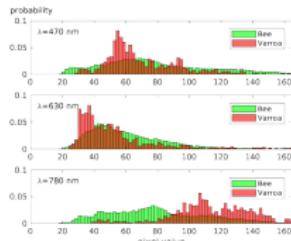


Figure 5: Histograms of bee and varroa pixel intensity values, for the spectral wavelengths 470 nm, 630 nm, and 780 nm respectively, recorded with the JAI camera. The image path is via the mirror-window-mirror, i.e., data were sampled with the setup given in figure 6. The image data for the histogram is the single bee with mite seen in figure 6.



Figure 6: The actual unprocessed camera view of the bees

spectively. The CM analysis was able to rank all wavelengths combinations, using one, two, three or four distinct wavelengths to give a ranking list of ‘best’ combination also taking the JAI camera spectrum into account.

The CM value of the actual choose wavelengths combination (470-630-780 nm) gave a rank just below the CM average score. This CM analysis was conducted after picking the actual used wavelengths, so later versions of the VMU might want to investigate a CM combination with a higher rank.

A specially designed diffuser and a number of narrow spectral LI were mounted in the camera focal diffuse illuminant fixtures.

Figure 6 disp along the passa era, with the gre with the NIR in

2.3.3. Real-time processing

A color and motion of 1296×96 from the camera over two separate sustained by ing frames real-

These data will post-process first matching rally coalescing producing a 24 the later image

Lossless real-time can be applied bandwidth that

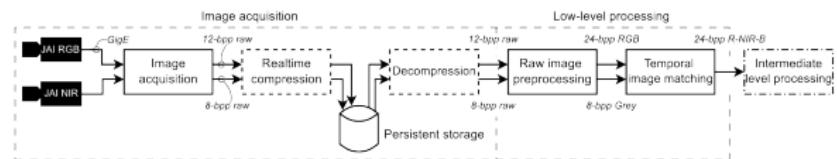


Figure 7: The low-level image processing pipeline. Raw camera images are stored on disk for later retrieval and post-processing. 12- and 8-bits per-pixel are used as the raw JAI/Bayer packed pixel format for the RGB and IR images respectively. Lossless, real-time compression can be introduced if persistent storage bandwidth is less than the raw-stream image rate of 93 MB/sec. The 12- and 8-bpp raw images from the network arrives out-of-order with respect to each other, hence the need for the temporal image matching.

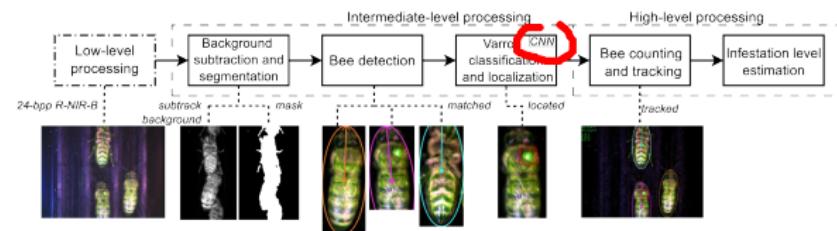
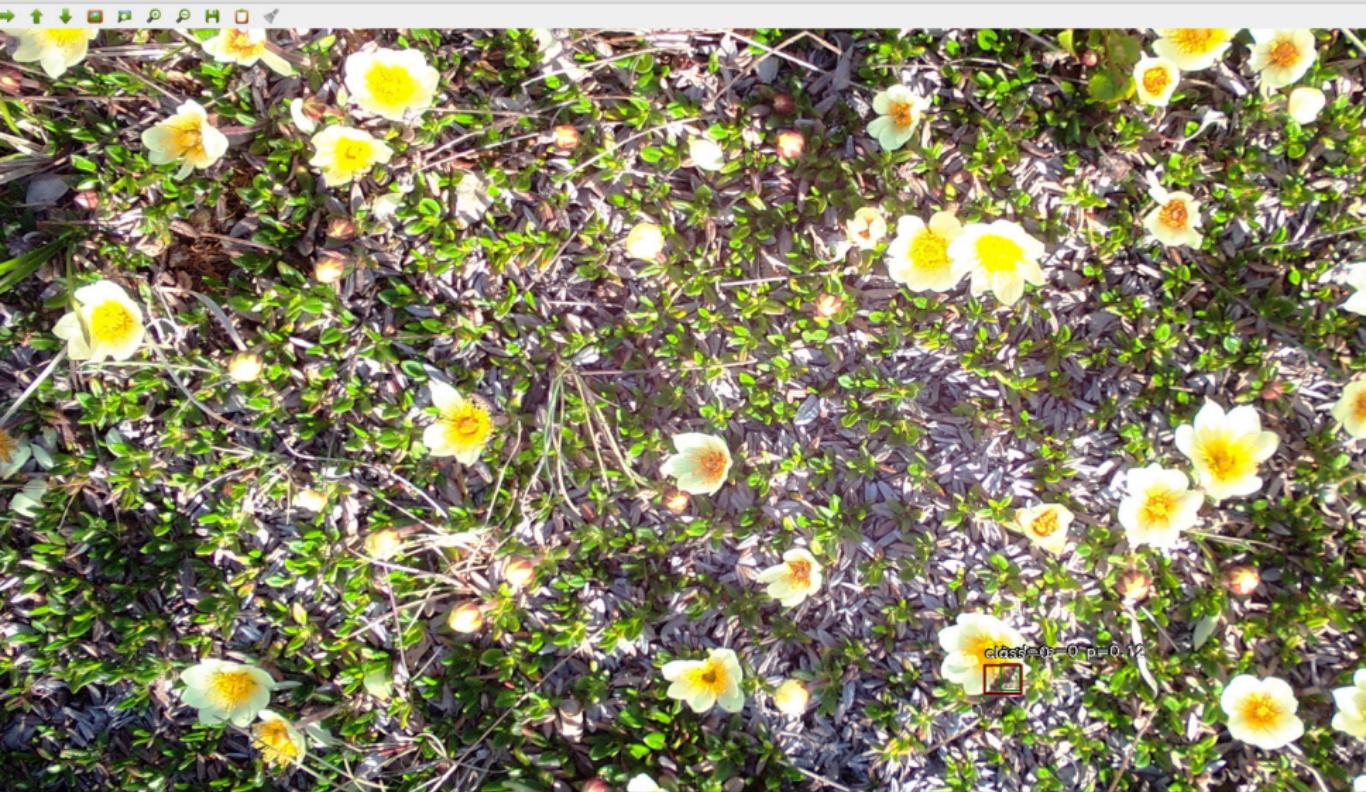


Figure 8: The processing pipeline of the intermediate- to high-level image processing algorithms to analyze and count the number of bees with *Varroa destructor*. A trained convolutional neural network (CNN) was used for the Varroa classification and localization stage.

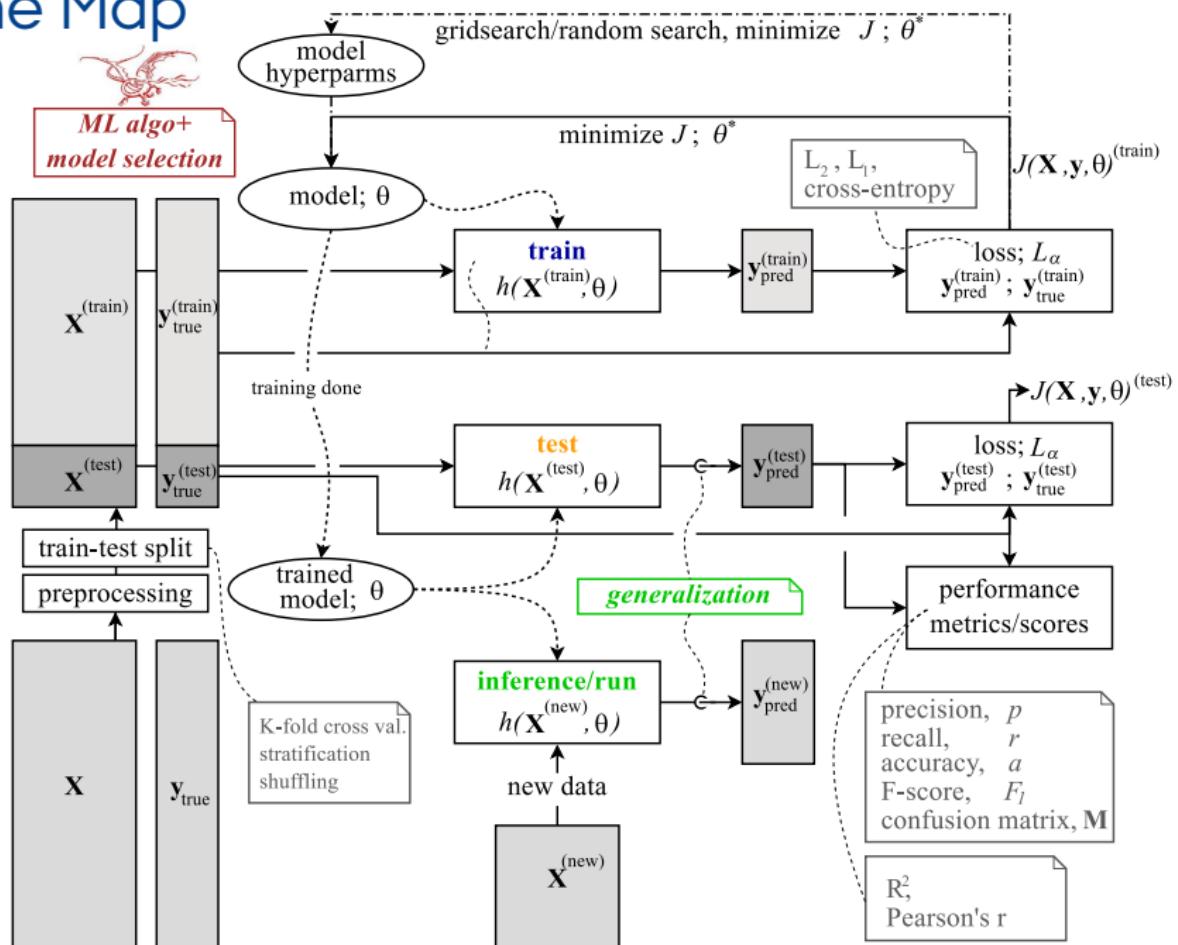
BA Project: generic annotation tool



MODEL CAPACITY



The Map



RESUMÉ: L02/performance_metrics.ipynb

Classification metrics

See the [Classification metrics](#) section of the user guide for further details.

<code>metrics.accuracy_score(y_true, y_pred[, ...])</code>	Accuracy classification score.
<code>metrics.auc(x, y[, reorder])</code>	Compute Area Under the Curve (AUC) using the trapezoidal rule
<code>metrics.average_precision_score(y_true, y_score)</code>	Compute average precision (AP) from prediction scores
<code>metrics.cohen_kappa_score(y1, y2[, labels, ...])</code>	Cohen's kappa: a statistic that measures inter-annotator agreement.
<code>metrics.confusion_matrix(y_true, y_pred[, ...])</code>	Compute confusion matrix to evaluate the accuracy of a classification
<code>metrics.f1_score(y_true, y_pred[, labels, ...])</code>	Compute the F1 score, also known as balanced F-score or F-measure
<code>metrics.log_loss(y_true, y_pred[, eps, ...])</code>	Log loss, aka logistic loss or cross-entropy loss.
<code>metrics.precision_score(y_true, y_pred[, ...])</code>	Compute the precision
<code>metrics.recall_score(y_true, y_pred[, ...])</code>	Compute the recall
<code>metrics.roc_auc_score(y_true, y_score[, ...])</code>	Compute Area Under the Receiver Operating Characteristic Curve (ROC AUC) from prediction scores.
<code>metrics.roc_curve(y_true, y_score[, ...])</code>	Compute Receiver operating characteristic (ROC)
<code>metrics.zero_one_loss(y_true, y_pred[, ...])</code>	Zero-one classification loss.

Regression metrics

See the [Regression metrics](#) section of the user guide for further details.

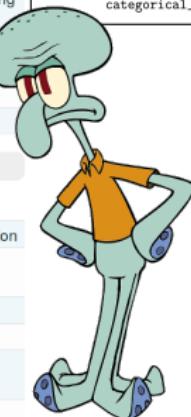
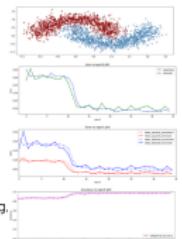
<code>metrics.explained_variance_score(y_true, y_pred)</code>	Explained variance regression score function
<code>metrics.max_error(y_true, y_pred)</code>	max_error metric calculates the maximum residual error.
<code>metrics.mean_absolute_error(y_true, y_pred)</code>	Mean absolute error regression loss
<code>metrics.mean_squared_error(y_true, y_pred[, ...])</code>	Mean squared error regression loss
<code>metrics.mean_squared_log_error(y_true, y_pred)</code>	Mean squared logarithmic error regression loss
<code>metrics.median_absolute_error(y_true, y_pred)</code>	Median absolute error regression loss
<code>metrics.r2_score(y_true, y_pred[, ...])</code>	R ² (coefficient of determination) regression

Notes on Keras MLPs

Typical Keras MLP Supervised Classifier setup:

- ▶ loss function
`loss='categorical_crossentropy'`
- ▶ metrics collected via history
`metrics=[`
 `'categorical_accuracy',`
 `'mean_squared_error',`
 `'mean_absolute_error']`
`]`
- ▶ input lay.: categorical encoding,
- ▶ output lay.: softmax function.

And notice that Keras do *not* provide metrics like
precision, recall, F1
but instead
`categorical_accuracy`, `binary_accuracy`



Model capacity

Exercise: `capacity_under_overfitting.ipynb`

Dummy and Paradox classifier:

capacity fixed ~ 0 , cannot generalize at all!

Linear regression for a polynomial model:

capacity \sim degree of the polynomial, x^n

Neural Network model:

capacity \propto number of neurons/layers

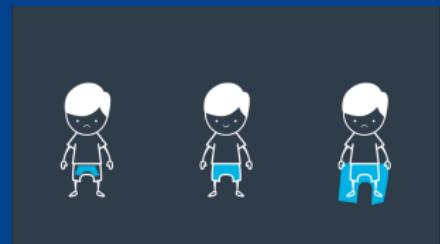
Homo sapiens ("modern humans"):

capacity \propto the IQ 'score' function?

⇒ **Capacity** can be hard to express as a quantity for some models, but you need to choose..

⇒ how to choose the **optimal capacity**?

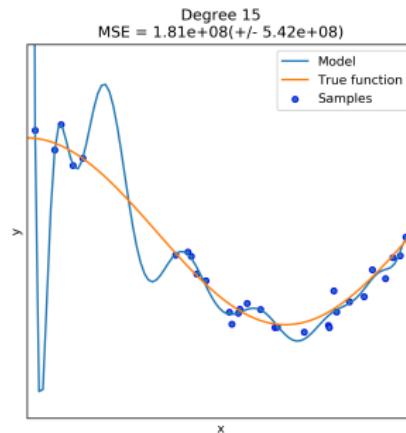
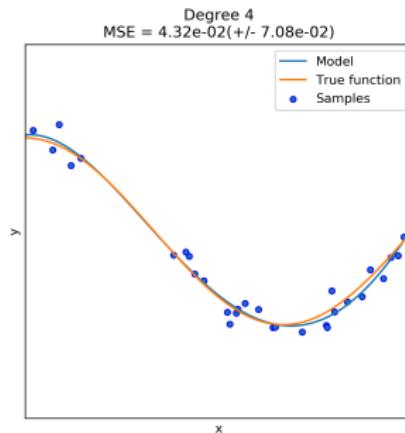
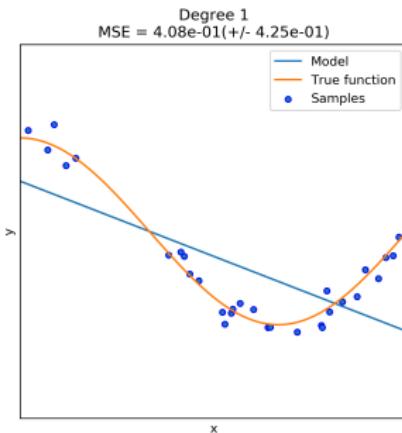
UNDER- AND OVERFITTING



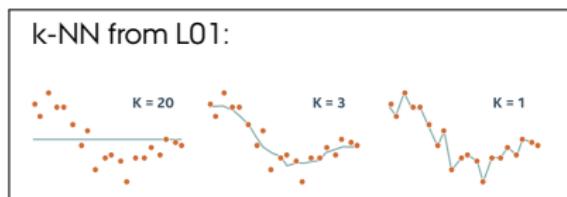
Under- and overfitting

Exercise: [capacity_under_overfitting.ipynb](#)

Polynomial linear reg. fit for underlying model: $\cos(x)$

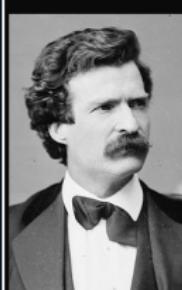


- ▶ underfitting:
capacity of model too low,
- ▶ overfitting:
capacity too high.



⇒ how to choose the **optimal capacity**?

GENERALIZATION ERROR



All generalizations are false, including this one.

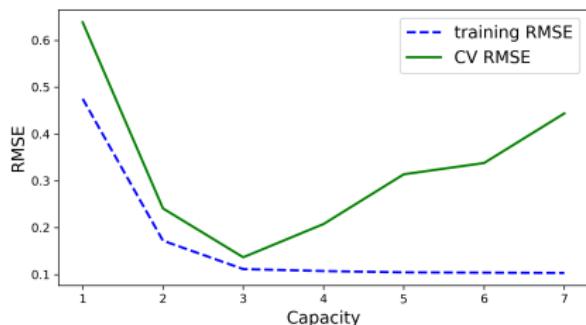
(Mark Twain)

Generalization Error

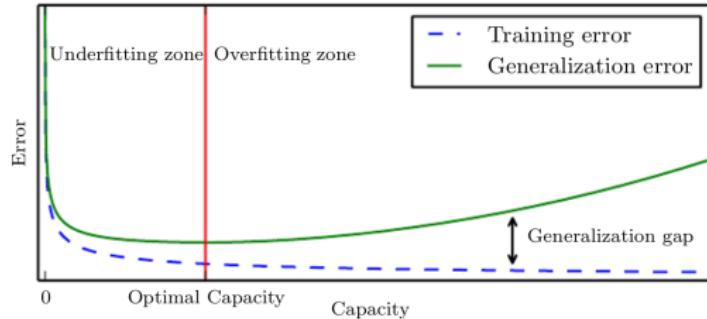
Exercise: generalization_error.ipynb

RMSE-capacity plot for lin. reg. with polynomial features

(capacity \sim degree of poly)



(Figure 5.3 from [DL])



Inspecting the plots from the exercise (.ipynb) and [DL], extracting the concepts:

- ▶ training/generalization error,
- ▶ generalization gap,
- ▶ underfit/overfit zone,
- ▶ optimal capacity (best-model, early stop),
- ▶ (and the two axes: x/capacity, y/error.)

Generalization Error

Definition of ML:

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .”

— Mitchell (1997).

Generalization Error

Exercise: generalization_error.ipynb

NOTE: three methods/plots:

- i) via **learning curves** as in [HOML],
- ii) via an **error-capacity** plot as in [GITHOML] and [DL],
- iii) via an **error-epoch** plot as in [GITHOML].

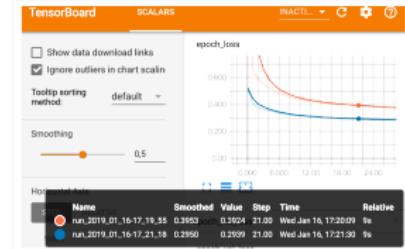
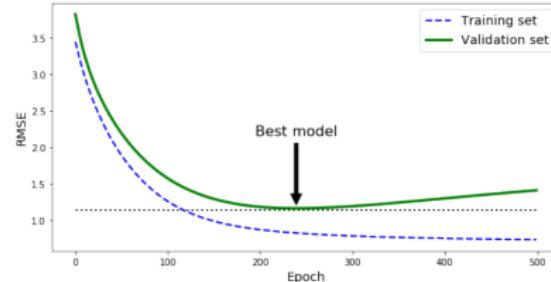
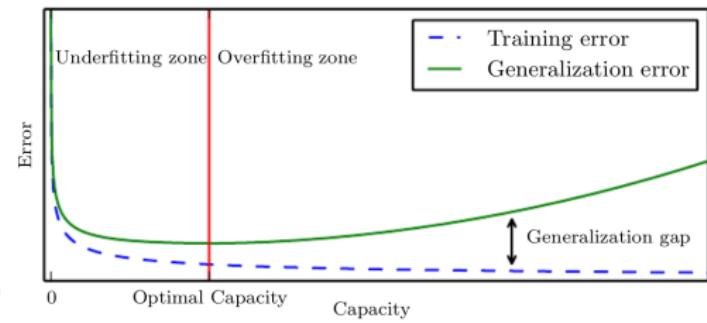
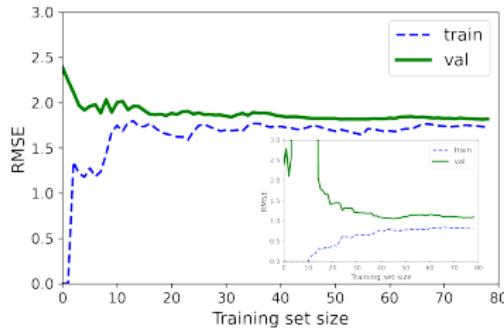


Figure 10-16. Visualizing Learning Curves with TensorBoard