



Artificiell Intelligens

Kurs introduktion



Agenda

- Kursmål
- Kursens innehåll
- Schema
- Hur

Kursmål

Kunskaper:

- Förstå vad artificiell intelligens och maskininlärning (ML) är.
- Förstå hur man utvecklar maskininlärningsmodeller.
- Förstå olika typer av algoritmer.

Färdigheter:

- Utveckla olika maskininlärningsmodeller.
- Använda relevanta pythonbibliotek.

Kompetenser:

- Resonera kring valet av ML algoritm.
- Utvärdera och optimera en modell.
- Redovisa resultat på ett professionellt sätt.

Kursens innehåll

1. Hur definieras Artificiell Intelligens (AI) och Maskininlärning?
2. Typer av AI med exempel på varje typ
 - a. Unsupervised – förstå/deskriptiv/reaktiv
 - b. Supervised -prediktera/proaktiv
 - c. Reinforcement learning – handling
 - d. Generative algorithms – kreativitet
 - e. Neural networks
3. Data cleaning & preparation
4. Model development
 - a. Träning och prediktion
 - b. Bias-variance trade-off
 - c. Model accuracy
 - d. Feature selection
 - e. Model selection
 - f. Model hyperparameter optimization
5. Markup language & reports
6. Applications of AI - Hur
 - a. Image analysis
 - b. Natural language processing
 - c. Virtual assistants
 - d. Predictive analytics, demand forecasting, maintenance
 - e. Robotics

Schema

AI	3	Intro & Applications of AI 1& 2	
	4	Regression, enkel linjär, multipel linjär, polynom(teori)	
	5	Programmera regression i Jupyter notebook	
	6	SVR, Decision trees, random forest, teori & Notebooks	
	7	R2 mått, Jämföra regressionsmodeller	
	8	Klassificering, Logistisk regression	
	9	Confusion matrix, accuracy, precision, recall	
	10	K-Nearest,SVM, SVM Kernel	
	11	Bayes teorem,Naive Bayes	
	12	Klassificering med Decision Tree och Random Forest.	
	13	Klustring intro	
	14	Artificiella Neurala nätverk, teoretisk genomgång	
	15	Regression och Klassificering, ANN i Tensorflow	
	16	Jämföra klassificeringar	
	17	Dela ut inlämningsuppgift, Reinforcement learning	
	18	Convolutional Neural Networks (CNN) och bildanalys	
	19	AWS	
	20	SageMaker Intro	
	21	Träna modell i cloud, Inlämning av uppgift	
	22	Köra tränad modell	
	23	Inlämning av URL till REST API, Avslutning	

Hur

Schema som vanligt om inte annat anges.

Måndagar 17:00 - 18:00 -Handledning

Onsdagar 16:00 - 19:00 - med raster

Kommer spelas in.

Övningsuppgifter och kompletterande material kommer ges löpande.

Examination

- Uppgift i slutet av kursen.

Extra info:

Kursens originalskapare : Bengt Holm

Baserad på Udemys kursen

"Machine Learning A-Z™:

Hands-On Python & R In Data Science"

Jag kommer lägga till mer.



Applications of AI

Part 1



Agenda

- Input
- Vad är AI?
- Types of AI
 - Some examples and discussions
- Fun examples
- Applications of AI
 - Natural language Processing
 - Image analysis
- Titanic - full example walkthrough

Input - Vår egen hjärna

- Människor har ett antal inputs till sin hjärna:
 - Hörsel
 - Syn
 - Lukt
 - Smak
 - Känsel
 - Humör
- Vad vi gör beror på vad vi får in via dessa signaler, oftast omedelbart utan att vi själva tänker på det.
- Personligt exempel: Ser jag damm på mitt tangentbord eller mus så blir min omedelbara output att blåsa bort dammet.



Input - Data

- Ett AI system måste få in data från en extern källa. Datan i sig kan vara vad som helst som kan omvandlas till data (1 och 0).
 - Bilder
 - Video
 - Ljud
 - Text
 - Rörelser
 - Röst

Vad är AI?

- Vad är intelligens? - Filosofisk fråga?
- Få maskiner att tänka själva? - Vad innebär det?
- Vilka områden av AI känner du till?
Vad kan du komma på? - 10 min
- Inom de nämnda områden- vilken sorts data tror du behövs?
(Vilken information behövs för att kunna nå målet?) - 15 min

Types of AI - Predictive Analysis

Predictive Analysis - likelihood of future events based on historical data

- Förutspå framtiden?
 - What will happen?
 - What if?
-
- Vilka applicationsområden faller under denna kategori? - 10 min

Types of AI - Predictive Analysis

Predictive Analysis

- Dating
- Trading
- Retail
- E-commerce

Types of AI - Image Analysis

“Extraction of meaningful information from images.”

- Wikipedia

- What is it?
 - What is needed for me to answer the above question?
 - Parts?
-
- Vilka applicationsområden faller under denna kategori? - 10 min

Types of AI - Image Analysis

Image Analysis

- Medicine
- Facial Recognition
- Cars
- Text
- Image enhancement

Types of AI - Classification

- För att träna ett AI system så måste man först ge den exempeldata och sedan testdata.
- Exempeldata: Jag har 300 bilder (100 kvadrater, 100 cirklar och 100 trianglar). Jag tar 20st bilder av varje form och matar det till AI systemet med en förklaring till systemet vilken form det är.
- Vad för exempel på klassifikationssystem kan ni komma på? - 10 min
Vad för information behövs för att avgöra detta?

Types of AI - Classification

Classification - known, 2 or more

- Insurance
- Loan
- Iris
- Medical diagnosis
- Character recognition
- Image
- Document sentiment

Types of AI - Clustering

Ett AI system kan också lära upp sig själv istället för med testdata

- Istället för att ha handmatad data som är förmarkerad, så kan ett AI system istället veta vad som är en positiv- och negativ outcome. Sedan måste AI Systemet själv hitta rätt input för att uppnå resultat.
- Detta är en långsammare inlärningsmetod, men i längden mycket mer effektiv då AI Systemet lär sig av sig själv och blir bättre med tiden.
- Kommer ni på något exempel där “klassen” inte är självklar?

Types of AI - Clustering

Clustering - unknown/less known

- Spam
- Markering
- Song / Film
- Text analysis - themes
- Fake news
- Recommendation engines
- Biology

Var används AI System

- AI system kan i princip användas var som helst, så länge den tar in data den kan bearbeta. Exempel redan nu är:
 - Bildidentifiering
 - Ansiktsidentifiering
 - Sjukvård
 - Schack (AlphaZero)
- AI systemen är beroende på den testdata som vi matar den med i början. Om inläringen är otillräcklig eller felaktig så kommer AI systemet att göra fel.

Ansiktsanimering

- Genom att använda sig av Keyframes (animations-term) och en bild på någon kan man via AI animera en persons ansikte till ett närmast naturligt exemplar.
 - Kan naturligtvis leda till farliga konsekvenser också



RESTORY...

Världsfred på Video



RESTOR\...



Del 2



RE STORY...

Applications of AI

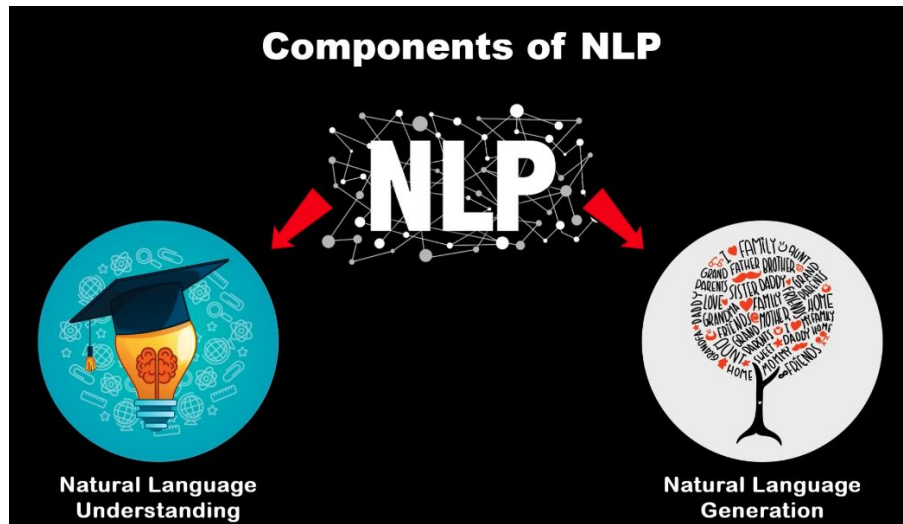
- Fraud and Risk Detection
- Healthcare
- Internet Search
- Targeted Advertising
- Website Recommendations
- Advanced Image Recognition
- Speech Recognition
- Airline Route Planning
- Gaming
- Augmented Reality

Machine Learning is the process of getting machines to automatically learn and improve from experience without being explicitly programmed.

Natural language processing

A subfield of linguistics, computer science, and AI.

Understanding and Generation
- which one is harder?



Difficulties

Not only are there hundreds of languages and **dialects**, but within each language is a unique set of **grammar** and **syntax** rules, **terms** and **slang**.

When we write, we often **misspell** or **abbreviate** words, or omit **punctuation**.

When we speak, we have **regional accents**, and we **mumble**, **stutter** and **mix** languages (loan words).

Ambiguity! - lexical, syntactic, referential
Inflection.

RE STORY...

Natural language processing

- **Tokenization** - breaking strings into parts
- **Stemming** - base/root form
- **Lemmatization** - maps words into common root
- **POS** tags - Parts Of Speech - noun, verb etc.
- **Named Entity Recognition** - Location, organization, person, thing.
- **Chunking** - grouping of words to get e.g. a known phrase “the pink panther”

nltk - python toolkit

- syntactic and semantic analysis

Syntax divides up sentences and uses things like *grammar rules* or *basic word forms* to understand a piece of text.

Semantics extracts the *meaning* behind it all. Using context, and tools like *word categorization*, or *meaning databases*, it discovers the *intention* behind using certain words.

Image analysis - steps

Image enhancement improves the quality of an image in order to extract hidden information from it for further processing.

Image restoration also improves the quality of an image, mostly by removing possible corruptions in order to get a cleaner version. This process is based mostly on probabilistic and mathematical models and can be used to get rid of blur, noise, missing pixels, camera misfocus, watermarks, and other corruptions that may negatively affect the training of a neural network.

Color image processing includes the processing of colored images and different color spaces.

Depending on the image type, we can talk about pseudocolor processing (when colors are assigned grayscale values) or RGB processing (for images acquired with a full-color sensor).

Image compression and decompression allow for changing the size and resolution of an image.

Compression - reducing the size and resolution.

Decompression - restoring an image to its original size and resolution.

These techniques are often used during the image augmentation process. When you lack data, you can extend your dataset with slightly augmented images. In this way, you can improve the way your neural network model generalizes data and make sure it provides high-quality results.

RESTORY...

Image analysis - steps

Morphological processing describes the shapes and structures of the objects in an image.

Morphological processing techniques can be used when creating datasets for training AI models. In particular, morphological analysis and processing can be applied at the annotation stage, when you describe what you want your AI model to detect or recognize.

Image recognition is the process of identifying specific features of particular objects in an image. AI-based image recognition often uses such techniques as object detection, object recognition, and segmentation.

Representation and description is the process of visualizing and describing processed data. AI systems are designed to work as efficiently as possible. The raw output of an AI system looks like an array of numbers and values that represent the information the AI model was trained to produce. Yet for the sake of system performance, a deep neural network usually doesn't include any output data representations. Using special visualization tools, you can turn these arrays of numbers into readable images suitable for further analysis.

Image analysis - group

1. Find examples of image analysis.
 2. Find an easy tutorial on kaggle and try to understand it. (You don't have to understand everything though!)
 3. Present.
 - a. What is the aim of the tutorial?
 - b. What is the data?
 - c. What transformations are performed on the data?
3. Present.
 - d. What EDA is performed on the data?
 - e. What are the difficulties in trying to understand it? Be specific!

Titanic example

[A simple example](#)

[Extensive example](#)

[Full of Transformation](#)



Scikit-learn

Mini intro



REPOSITORY...

Scikit-learn

Dataset loading utility

- **loader** - small datasets, e.g. Iris
- **fetcher** - larger datasets, classes or samples
- **generated** - random sample generators that can be used to build artificial datasets of controlled size and complexity. **make**
- **fetch_openml** - openml.org is a public repository for machine learning data and experiments, that allows everybody to upload open datasets.

Preprocessing package - provides several common utility functions and transformer classes to change raw feature vectors into a representation that is more suitable for the downstream estimators.

- **StandardScaler** - centered around zero and have unit variance.
- **MinMaxScaler** - [0, 1].
- **MaxAbsScaler** - [-1, 1], sparse data (many values are zero).
- **RobustScaler** - many outliers.
- **Normalize** - unit norm, independent of the distribution of the samples.
- **OrdinalEncoder** - converts categorical features to integer codes.
- **KBinsDiscretizer** - binning, partitions continuous features into discrete values.
- **Binarizer** - feature binarization: thresholding numerical features to get boolean values.

Scikit-learn

Pipeline - chain of transforms (and estimators).

model selection - different **splits** (and more)

- train_test_split
- KFold
- Cross validation

feature selection - dimensionality reduction on sample sets, improve accuracy scores or boost performance on very high-dimensional datasets.

- **VarianceThreshold** - removes all features whose variance doesn't meet some threshold.

feature extraction - extracts features from raw data (text and images).

Imputation - scikit-learn estimators assume that all values in an array are numerical, and that all have and hold meaning.

- **SimpleImputer** - mean or most frequent if categorical
- **IterativeImputer** - models each feature with missing values as a function of other features, and uses that estimate for imputation.
- **KNNImputer** - fills in missing values using the k-Nearest Neighbors approach

Övning

Använd dig av följande scikit-learn funktioner på Titanic dataset:et.

- `fetch_openml`
- `OrdinalEncoder`
- `IterativeImputer`
- `KBinsDiscretizer`
- `StandardScaler`

[Scikit-learn example on youtube](#)

Sammanfattning

- Input = data
- Vad är AI?
- Types of AI - some examples
 - Predictive Analysis
 - Image Analysis
 - Classification
 - Clustering
- Ansiktsanimering - exempel
- Applications of AI
 - Natural language Processing
 - Image analysis - egen analys
- Titanic - full example walkthrough

Länkar

- [GTP-3 intervjuv](#)
- [GTP-3 intervjuv 2](#)
- [Brittiskt program om AI med Gemma Chan från programmet Humans som blir "kopierad".](#)
- [Data science project suggestions](#)
- [Scaling comparisons](#)
- [NLP example](#)
- [Text generering](#)
- [Klassificering av kläder](#) (tensorflow)
- [Ett till Titanic exempel](#)
- [Ted talk om bias](#)
- [Scikit-learn LabelEncoder](#)