

Datadriven analys och maskinintelligens

(Data science and machine intelligence)

**Profil på Y-programmet
HT-2024**

Erik G. Larsson, ISY
erik.g.larsson@liu.se, 013-281312

Vad är
"data science"?

\$600 to buy a disk drive that can store all of the world's music

5 billion mobile phones in use in 2010

30 billion pieces of content shared on Facebook every month

40% projected growth in global data generated per year vs. **5%** growth in global IT spending



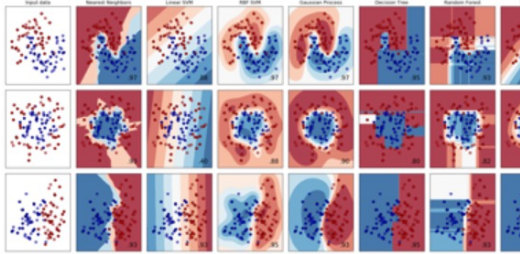
Vad är maskin-intelligens?

Classification

Identifying which category an object belongs to.

Applications: Spam detection, image recognition.

Algorithms: SVM, nearest neighbors, random forest, and more...

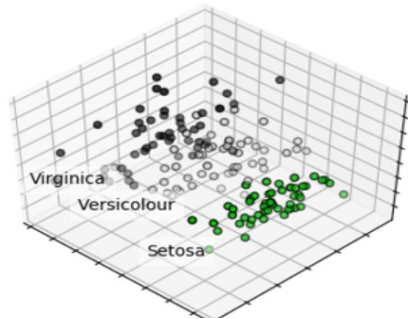


Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency

Algorithms: k-Means, feature selection, non-negative matrix factorization, and more...

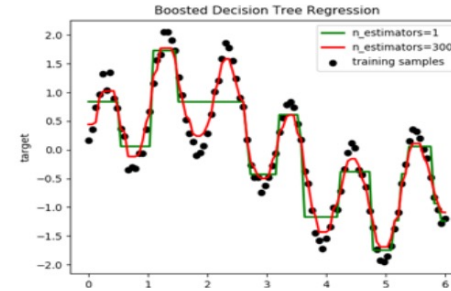


Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.

Algorithms: SVR, nearest neighbors, random forest, and more...

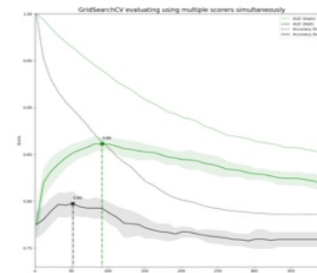


Model selection

Comparing, validating and choosing parameters and models.

Applications: Improved accuracy via parameter tuning

Algorithms: grid search, cross validation, metrics, and more...



Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes

Algorithms: k-Means, spectral clustering, mean-shift, and more...

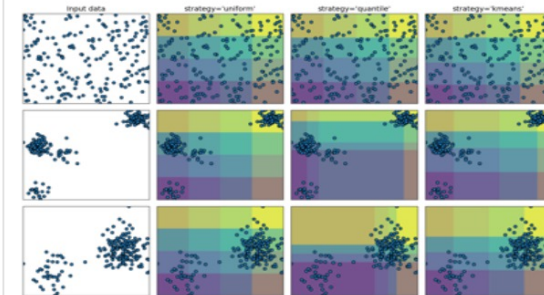


Preprocessing

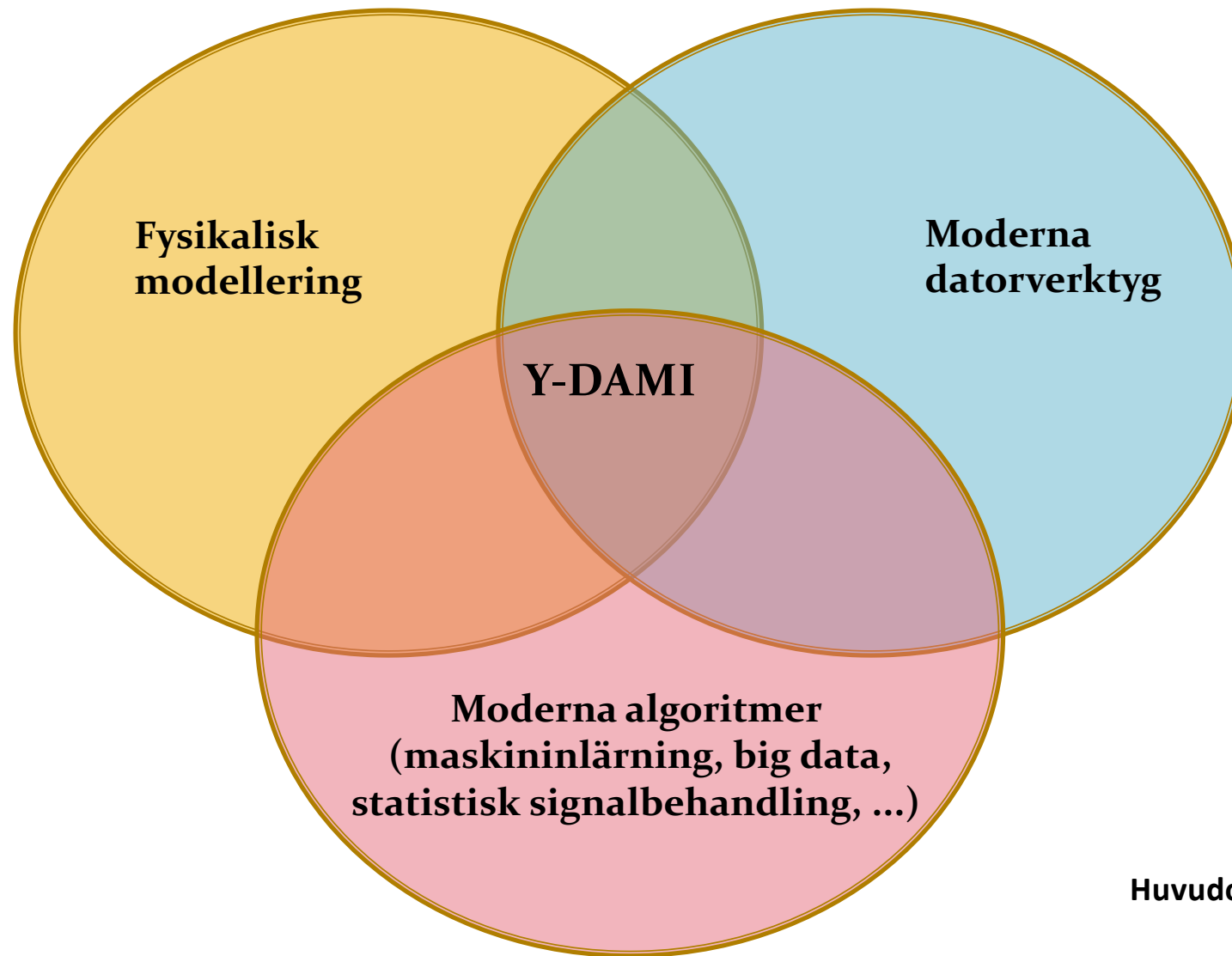
Feature extraction and normalization.

Applications: Transforming input data such as text for use with machine learning algorithms.

Algorithms: preprocessing, feature extraction, and more...



Y-profilen “datadriven analys och maskinintelligens” (Y-DAMI)



Huvudområde: datateknik

6 obligatoriska kurser + 1 profilkurs + CDIO (+ 6 helt valfria)

ÅK	Period	Block 1	Block 2	Block 3	Block 4
4HT	7ht1	Maskininläring	Multidimensionell signalanalys (1)	Modellering och inläring för dynamiska system	Detektion och estimering av signaler
	7ht2		Komplexa nätverk och stora datamängder	Multidimensionell signalanalys (2)	
4VT	8vt1	Algoritmisk problemlösning (1)	Neuronnät och lärande system	Data Mining - Clustering and Association Analysis	
		Numerisk linjär algebra	Språkteknologi		
			Avancerad programmering (1) i C++		
	8vt2	Optimeringslära fortsättningskurs	Bayesianska metoder	Big data analytics	
Avancerad programmering (2) i C++		Programmering av paralleldatorer - metoder och verktyg	Sensorfusion		
5HT	9ht1	Signalteori		Artificiell intelligens	Multivariat statistik*
		Avancerad maskininläring		Optimering av stora system	Digital bildbehandling grundkurs
		Modern kanalkodning, inferens och inläring			Projekturs (TSKS23 eller TSBB11)
	9ht2	Medicinsk bildanalys	Text mining	Vetenskaplig metod	Projekturs (TSKS23 eller TSBB11)
		Databasteknik	Digital signalbehandling		

Obligatorisk kurs	Valbar kurs
-------------------	-------------

* = ges vartannat år

TSRT92 Modelling and Learning for Dynamical Systems

Exempel på innehåll:

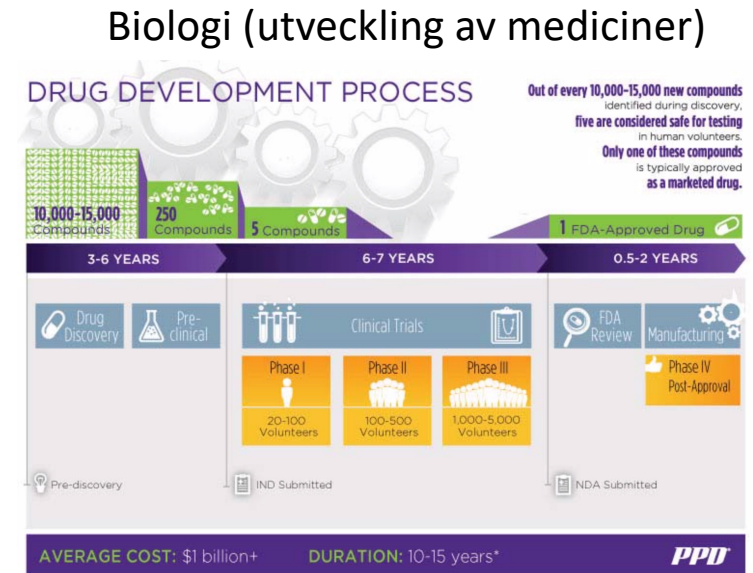
- Modellbygge utifrån fysikaliska samband.
- Differential- och differensekvationer.
- Differentialalgebraiska (DAE-) modeller.
- Blackboxmodeller. Neurala nätverk.
- Identifiering: Frekvens och transientanalys.
- Korrelations- och spektralanalys.



Flygplansprestanda



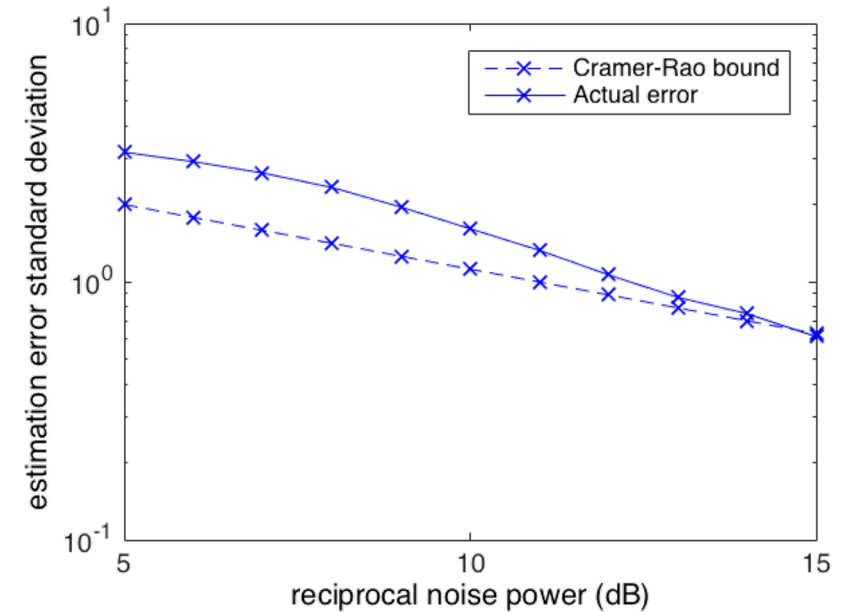
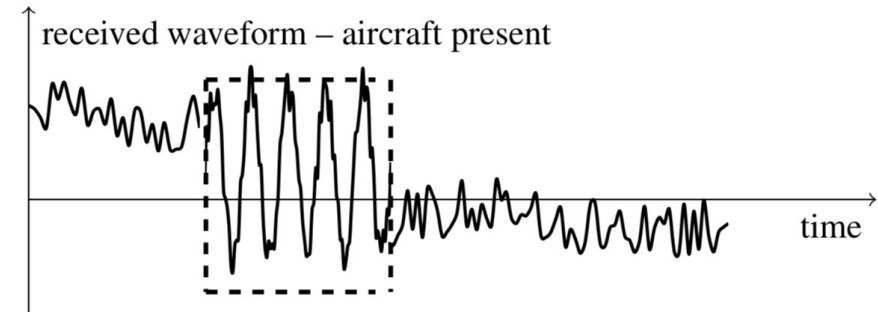
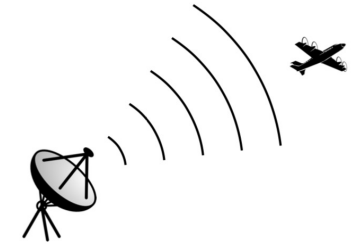
Process-
industri



TSKS15 Detection and Estimation of Signals

- **Detektion** av signaler i brus, exempel på tillämpningar:
 - detektion av radarekon
 - vilken vågform skickades?
 - mönster i dataströmmar
- **Estimering** av signalparametrar, tillämpningar:
 - avstånd och riktning till ett radarmål
 - orientering av en kamera
 - identifiering av okända system, t.ex. LTI-impulssvar
- **Prestandamått och fundamentala gränser**

Praktiska algoritmer + byggblock i stora system för dataanalys och maskininlärning

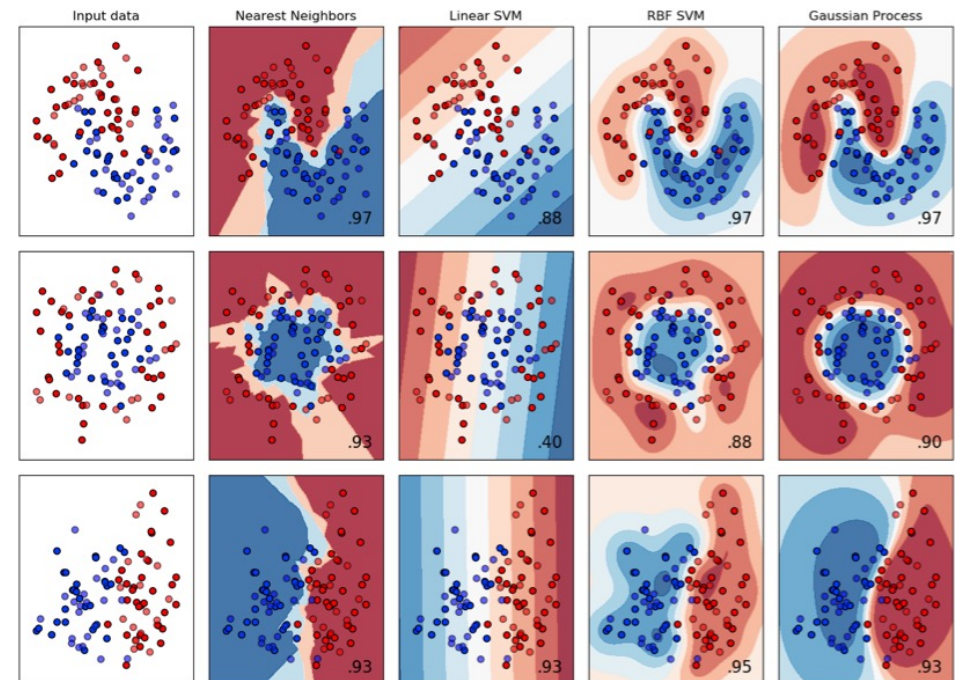


TDDE01 Machine Learning

Introduktion till maskininlärning, med speciellt fokus på regressions- och klassificeringsproblem.

Exempel på innehåll:

- Översikt av maskininlärningsområdet + tillämpningar
- Unsupervised och supervised learning
- Prediktion. Klassificering.
- Nearest neighbors. Naïve Bayes. Diskriminantanalys. Korsvalidering. Modellval.
- Överanpassning. Bootstrap. Regression. Regularisering.
- Ridge regression. Lasso.
- Dimensionsreducering. PCA. ICA.
- Support vector machines. Beslutsträd.
- Gaussiska processer. Mixture models.



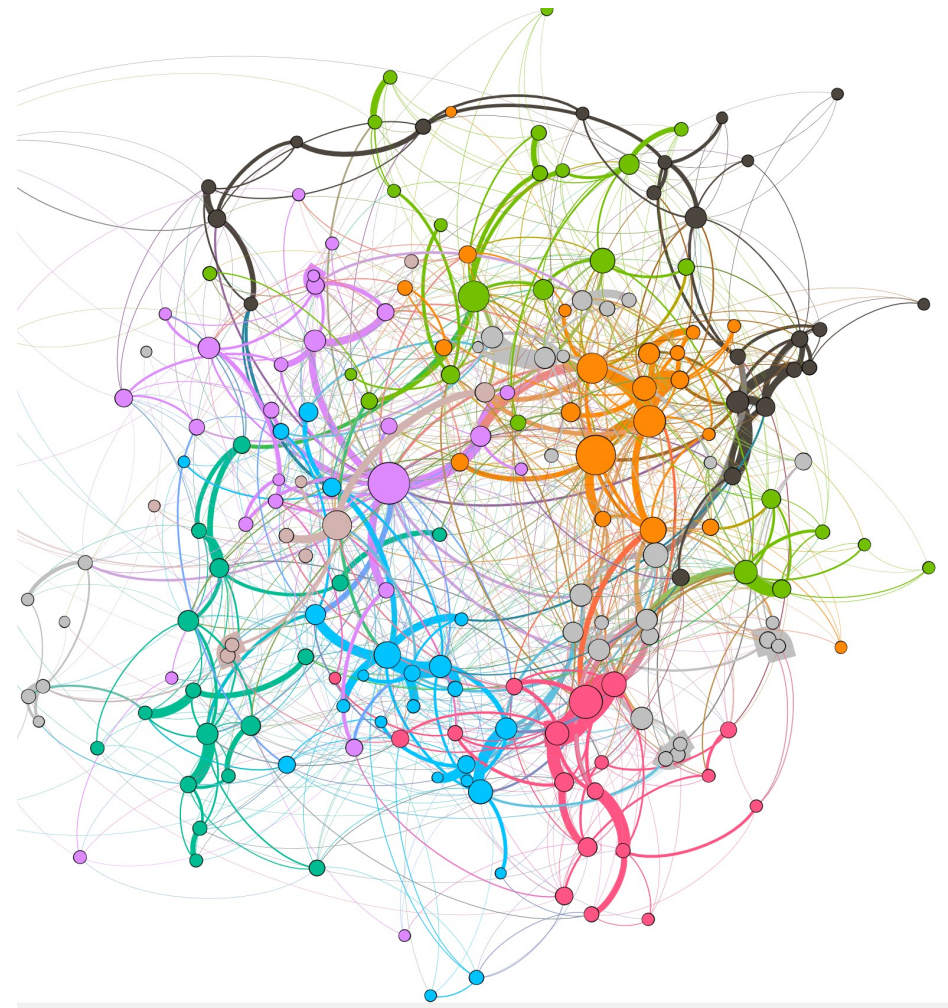
TSKS33 Complex Networks and Big Data

Komplexa nätverk beskriver hur objekt är sammankopplade, till exempel:

- routrar på Internet
- webbsidor
- människor i sociala nätverk
- städer i transportnätverk
- substanser i biologiska nätverk

Exempel på innehåll:

- Matematiska modeller för nätverk
- Metriker, t.ex. centralitet, assortativitet
- Sampling på nätverk
- Algoritmer för partitionering och “community detection”
- Slumpmodeller för nätverk
- Världen-är-liten-fenomen
- Graph learning

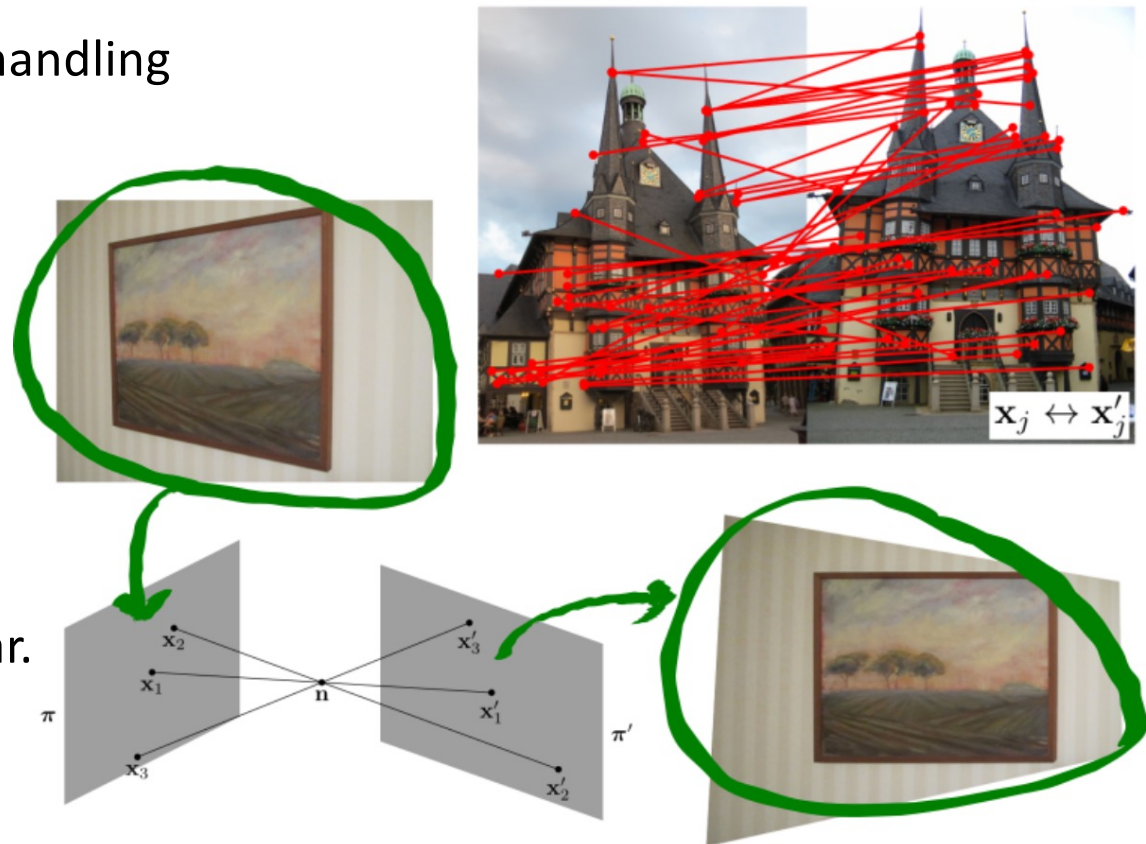


TSBB06 Multidimensionell Signalanalys

- Grundkunskaper inom signal- och bildbehandling
- Bakgrund för att gå TSBB11 (CDIO)

Exempel på innehåll

- Signalrum och signalbaser, dualbaser.
- Minstakvadratproblem
- Principalkomponentanalys.
- Ramer, wavelet- transform och filterbankar.
- Projektiva rum, homogena koordinater, homografier, kameraprojektioner, epipolargeometri.
- Representation och skattning geometriska objekt.

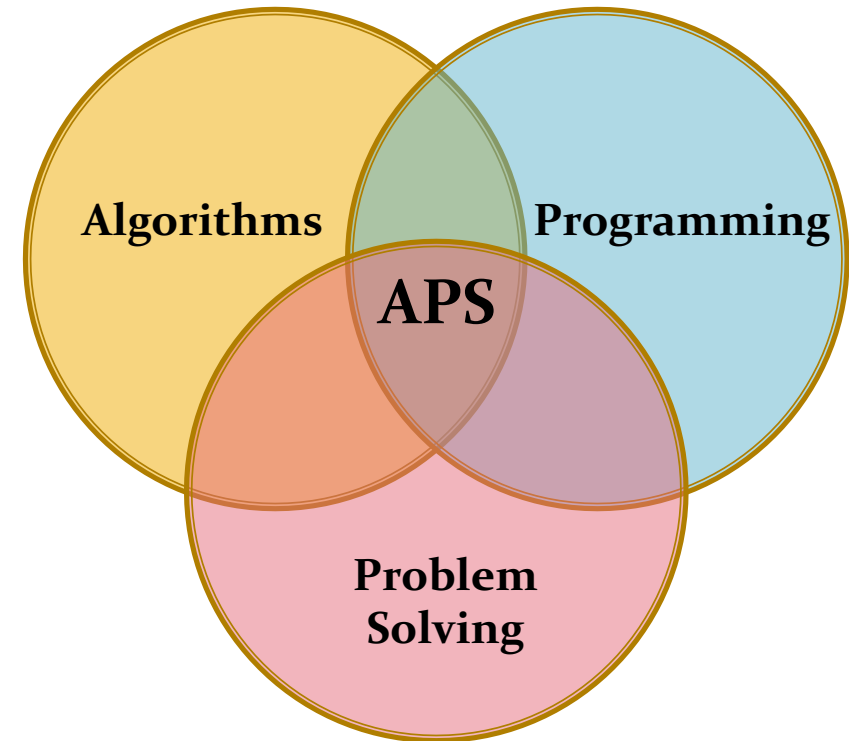


TDDD95 Algorithmic Problem Solving

- Algoritmer, datastrukturer och komplexitetsteori
- Analysera problem, välja eller konstruera en algoritm, analysera, implementera och verifiera
- Omfattande **labbserie** med programmeringsuppdrag

Labbarna handlar bland annat om:

- optimeringsalgoritmer
- ekvationssystem
- grafer
- strängar
- aritmetik
- geometri



CDIO-kurs:

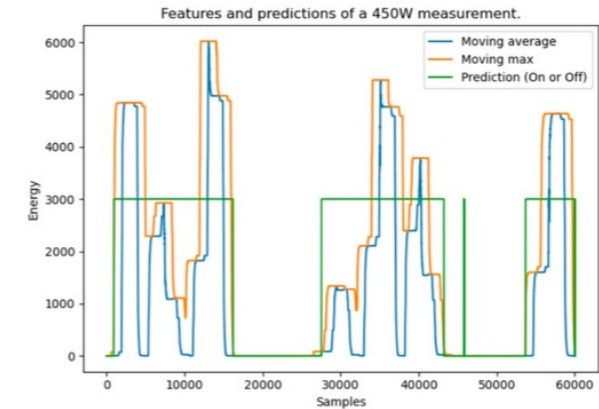
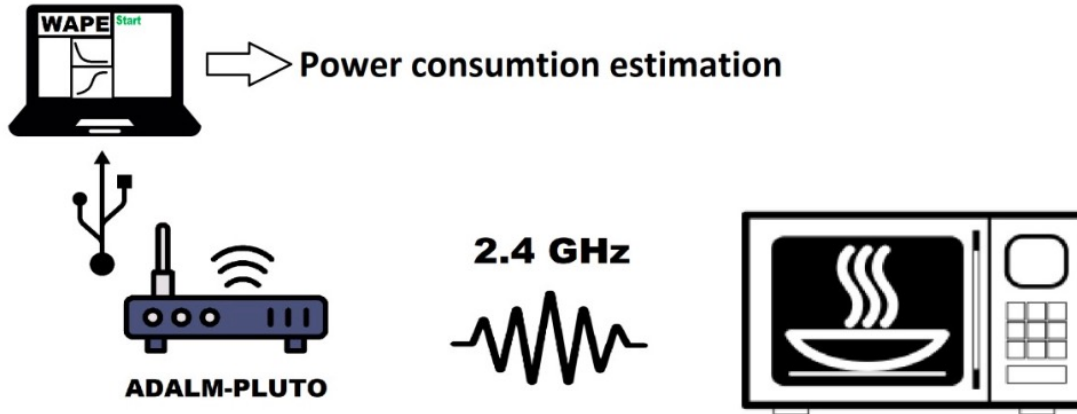
**TSKS23 (signalbehandling, kommunikation och nätverk), eller
TSBS11 (Bilder och grafik)**

Det finns två projektkurser att välja mellan: TSBB11 och TSKS23. De obligatoriska profilkurserna i åk 4 ger de förkunskaper som behövs.

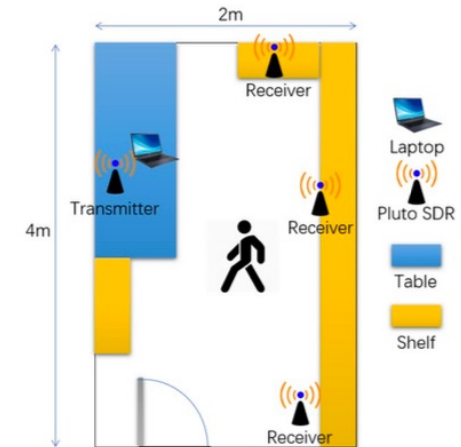
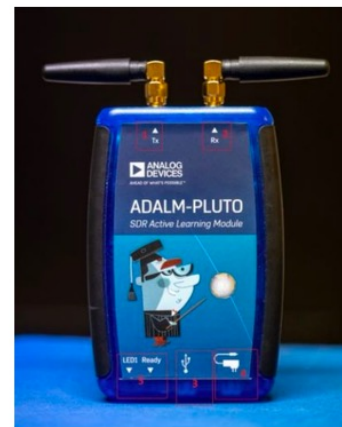
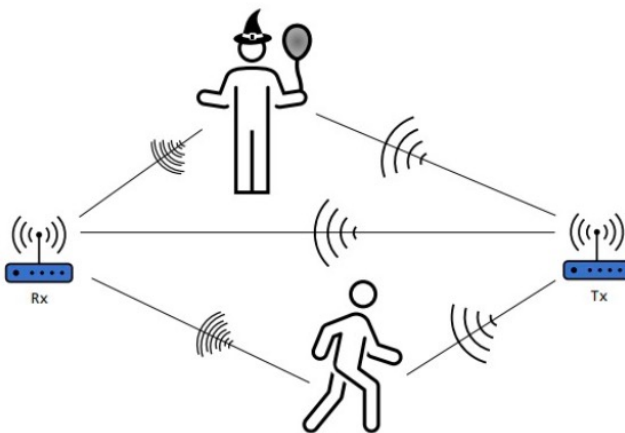
För er som vill välja TSBB11, är det önskvärt (men inte tvunget) att även läsa någon av kurserna TSBB08, TSBB09, TSBK07, TBMI26, TSBB15 eller TSBB19. (Flera av dessa är valbara profilkurser, och de andra kan väljas från det helt valfria kursutbudet på Y.)

TSKS23 - exempel på projekt från tidigare år

Energy consumption of microwave oven devices



Detection of static and dynamic indoor environments

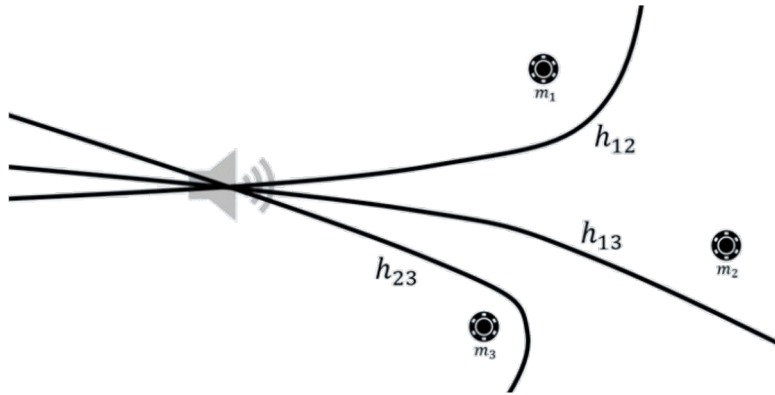


TSKS23 - exempel på projekt från tidigare år

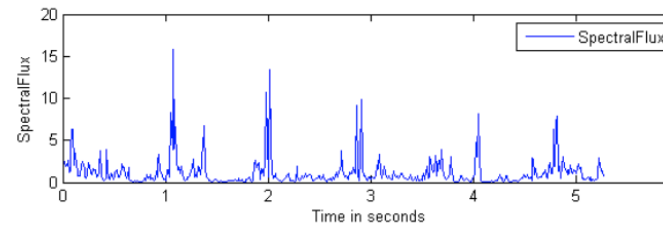
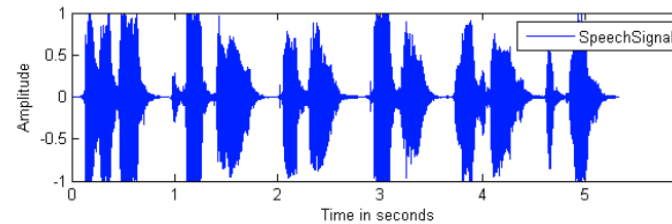
Lokalisering av ljudkällor



6-mic array



Transkription av musik



Exempel på exjobb, studenter från Y-DAMI tidigare år

- *Estimating Forest Variables from LiDAR Pointcloud Data Using a Deep Learning Approach*, Digital Edge, Stockholm
- *The Effect of Distortions Induced by Adaptive Antenna Arrays in GNSS Applications*, FOI, Linköping
- *Fault-Tolerant Over-the-Air Federated Learning with Clustered Aggregation*, LiU/ISY
- *Examining Difficulties in Weed Detection*, RISE
- *Finding Corresponding Regions In Different Mammography Projections Using Convolutional Neural Networks*, Sectra

Kontakt:

**Erik G. Larsson, profilansvarig
Linköpings universitet
Institutionen för systemteknik (ISY)**

erik.g.larsson@liu.se

013-281312