

## 2- SUPERVISED

- CLASS HYPOTHESES  $\rightarrow$  GENERALIZATION  $\rightarrow$  MARGIN  
     $\hookrightarrow$  VERSION SPACE
- VC-DIMENSION
- PAC LEARNING
- NOISE  $\begin{cases} \rightarrow$  TIPOLOGIE \\ \quad  $\hookrightarrow$  SOURCE \end{cases} \rightarrow
- APPRENDIMENTO  $\begin{cases} \text{SINGLE CLASS} \\ \text{MULTI CLASS} \rightarrow \text{1/K} \\ \text{REGRESSION} \end{cases}$
- MODEL SELECTION  $\begin{cases} \rightarrow$  BIAS VARIANCE \\ \quad  $\hookrightarrow$  UNDER/ OVER FITTING \\ \quad  $\hookrightarrow$  TRIPLE TRADE OFF \\ \quad  $\hookrightarrow$  CROSS VALIDATION \end{cases}

## 3- BAYESIAN DECISION THEORY

- BAYES' CLASSIFIER
- RISCHIO ATTESO AZIONI  $\rightarrow$  AZIONE REPERTO
- REGIONI DECISIONE
- REGOLE ASSOCIAZIONE  
     $\begin{cases} \text{SUPPORTO CONFIDENZA} \\ \text{LIFT} \\ \text{APRIORI} \end{cases}$

## 4 - PARAMETRIC METHODS

- STIMA PARAMETRI  $\rightarrow$  MLE

$\left\{ \begin{array}{l} \text{BERNOLLI} \\ \text{MULTINOMIALE} \\ \text{GAUSSIAN} \end{array} \right.$

- BIAS-VARIANCE  $\left\{ \begin{array}{l} \rightarrow \text{UNBIASED ESTIMATOR} \\ \rightarrow \text{CONSISTENT ESTIMATOR} \\ \rightarrow \text{ASYMPTOTICALLY UNBIASED} \end{array} \right.$

- BAYES ESTIMATOR  $\left\{ \begin{array}{l} \rightarrow \text{INTEGRAL (FULL)} \\ \rightarrow \text{MAP} \rightarrow \text{REL MLE} \\ \rightarrow \text{BAYES ESTIMATOR} \end{array} \right.$

- PARAMETRIC CLASSIFICATION

- 1) ASSUMO FORMA  $P(x|c_i)$
- 2) DEFINISCO  $g_i(x) = P(x|c_i)P(c_i)$
- 3) STIMO  $\theta$
- 4) APPLICO  $\theta$  a  $g_i(x)$

- POSSO FARE ASSUNZIONI SU PARAMETRI E SEMPLIFICARE

- THRESHOLD DECISIONE + RISCHIO a REJECT

- REGRESSIONE

- 1) Dato  $(x^t, r^t)$   $P(x, r) \sim P(r|x)P(x)$

- 2)  $r = f(x) + \varepsilon$   $\varepsilon \sim \mathcal{N}(0, \sigma^2)$

- 3)  $g(x|\theta)$  regressore su  $r$

- 4)  $P(r|x) \sim \mathcal{N}(g(x|\theta), \sigma^2)$

- 5) Calcolo  $L(\theta|x) \rightarrow E(\theta|x)$  (SSE)

- MINIMI QUADRATI

- REGRESSIONE LINEARE

- REGRESSIONE POLINOMIALE

- VALUTAZIONE  $\left\{ \begin{array}{l} \text{MSE} \\ \text{RMSE} \\ \text{R}^2 \end{array} \right.$

- BIAS VARIANCE TUNING

- ERRORE NON RIDUCIBILE + ERRORE

- MEDIA SU DATASET

$\rightarrow$  BIAS + VARIANCE

- MODEL SELECTION

- 1) CROSS VALIDATION
- 2) REGULARIZATION
- 3) OPTIMISM TERM
- 4) SRM
- 5) MDL
- 6) BAYESIAN MODEL SELECTION

# S-MODELLI MULTIVARIATI

## • PARAMETER ESTIMATION

Per  $\mu \in \mathbb{R}^d$  e  $\Sigma \in \mathbb{R}^{d \times d}$

↳ Produco stime biased

## • MISSING IMPUTATION $\begin{cases} \text{MEAN} \\ \text{REGRESSION} \end{cases}$

## • NORMALE MULTIVARIATA $\begin{cases} \text{DISTANZA MANHATTAN O} \\ \text{USO } \Sigma^{-1} \rightarrow \text{STD UNIT VAR} \\ \text{2-NORMALIZATION} \end{cases}$

- Interpretazione  $\mu$  e  $\Sigma$  e  $p$  (in  $\text{dim} = 2$ )

① MATRICE FULL

② MATRICE DIAGONALE

- Prevedere minima errore  $W^T x$

## • CLASSIFICAZIONE MULTIVARIATA

### - DISCRIMINANTE QUADRATICO

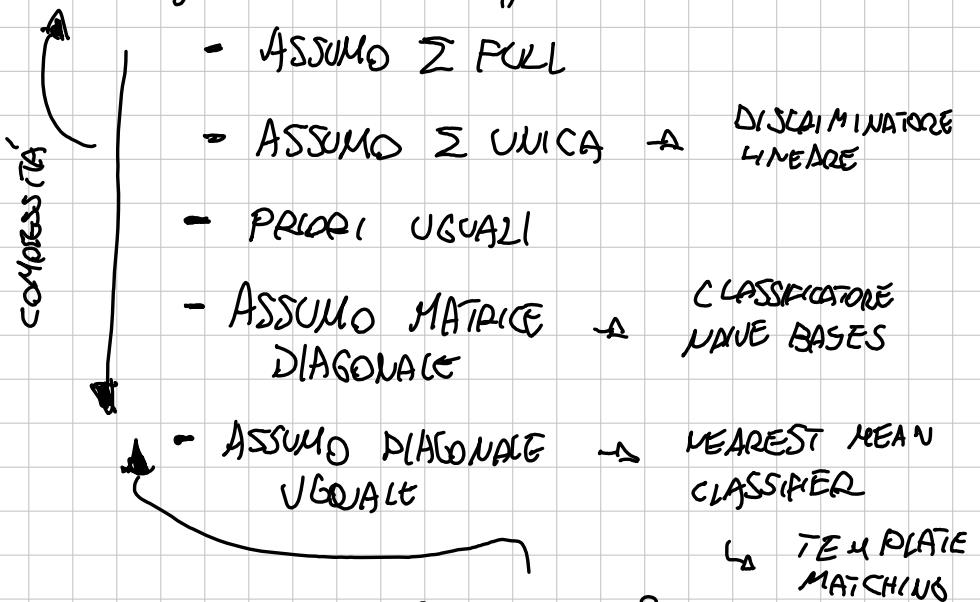
1) Assumo  $p(x|c_i) \sim \mathcal{N}(\mu_i, \Sigma_i)$

2) Assumo sample dove da single gruppo

3)  $g_i(x) = \log P(x|c_i) + \log(P(c_i))$

4) Stimare  $\mu_i, \Sigma_i$  (MLE me dare)

5) Plug stime in  $g_i(x)$



## • TUNING COMPLEXITY: QUANTO USARE?

↳ RDA

## • FEATURE DISCRETE

- BINARIE BERNOULLI  $p_{is} = P(x_s = 1 | c_i)$

- MULTINOMIALE  $p_{ijk} = P(x_s = v_k | c_i)$

## • REGRESSIONE

↳ MINIMIZZO ERRORE (LINEARE)  
SSE

$\rightarrow$  DERIVATA  $\rightarrow$  EQUAZIONI NORMALI

- MULTIVARIATE POLYNOMIAL REGRESSION

- BASIS FUNCTION





## 7 - CLUSTERING

- APPROCCIO SEMI PARAMETRICO

- MIXTURE DENSITIES [ COMPONENT PROPORTIONS ]

- K-MEANS

- DEFINISCO DEI REFERENCE VECTOR → MINIMIZZANDO ↓
- RICOSTRUISCO USANDO QUESTI → ERRORE RICOSTRUZIONE
- PROCEDURA ITERATIVA

CIRCOLO

0) REF VECTOR RANDOMIC

1) CACCIO ASSEGNAZIONE LABEL

2) MINIMIZZO ERRORE RICOSTRUZIONE

3) NUOVI REF VECT SARANNO MEDIA ELEMENTI: STESSA CLASSE

4) FERMO SE STABILIZZO

- LEADER CLUSTER

- SE ISTANZA MOLTO LONTANA CREA UN NUOVO CLUSTER
- CLUSTER CON TROPPE ISTANZE SPLITATO

- EXPECTATION MAXIMIZATION

- CERCO PARAMETRI COMPONENTI MASSIMIZZANDO LIKELIHOOD
- PROCEDURA ITERATIVA

EM

E: STIMO PARAMETRI NON OSSERVABILI DATE COMPONENT

M: UPDATE COMPONENT CON NUOVI PARAMETRI

E {  
1) VARIABLE INDICATOR ASSEGNAMENTO → MULTINOMIALE  
2) CALCOLO LIKELIHOOD PARAMETRI  
    ↳  $E[INDICATOR] = \text{POSTERIORI CLASSE}$   
3) OTTENGO SOFT LABEL ASSIGNMENT } stima

M {  
4) MASSIMIZZO L CON UPDATE PARAMS  
    ↳ LAGRANGE VINCOLATO → DERIV PARZ

ASSUNZIONI  
SEMPLICI

$P(x \in G_i)$

NORMALE COVARIANZA COMUNE

NORMALE COVARIANZA DIA GONALE

- MIXTURE OF LATENT VARIABLES

1) FACTOR ANALYSIS NEI CLUSTER

2) E STEP CON NORMALE CON FACTOR LOADINGS

3) M STEP STIMO FACTOR LOADING

- MIXTURE OF PROBABILISTIC PCA

• Mixture clustering con ← RIDURRE DIM  
← KNOWLEDGE EXTRA  
← RAPPRESENTAZIONE ← DISTRIBUITA  
← LOCALE

- SPECTRAL CLUSTERING

1) APPLICO LAPLACIAN EIGENMAP → VA NORMALIZZATA

2) APPLICO K-MEANS

- HIERARCHICAL

- NON PROBABILISTICO USE DISTANZE

- TIPI ← AGGLOMERATIVO  
← DIVISIVO

ELONGATED

- ACCORDAMENTO DISTANZA CLUSTER ← SINGLE LINK - CHAINING  
← COMPLETE LINK - COMPACT

- DENDROGRAM

- Ottimizzare K con ELBOW METHOD

## 8 - NON PARAMETRIC

### DENSITY ESTIMATION

- HISTOGRAM ESTIMATOR

- DIVIDO SPATIO IN BINS DATO CENTRO E LARGHEZZA

- NAIVE ESTIMATOR

- TRATTO X COME CENTRO E FISSO RAGGIO  $H/2$

- KERNEL ESTIMATOR (PARZEN WINDOWS)

- USO GAUSSIAN KERNEL PER SMOOTH ESTIMATE
- POSSO USARE MAHALANOBIS SE HO UN'ALTRA CURVA DI KERNEL

- K-NEAREST NEIGHBORS

- ORDINO IN MODO CRESCENTE DISTANZA DA ALTO, ELEN
- FISSO NUMERO ELEM CHE DEVONO ANDARE E CALCOLO LARGHEZZA BIN
- NON CONTINUA
- POSSO USARE KERNEL PER SMOOTH

### CLASSIFICATION

- KERNEL CLASSIFIER

- USO STIMATORE DELLE  $P(x|C_i)$
- STIMA MLE DI  $P(C_i)$

- KNN CLASSIFIER

- ASSEGNA CLASSE CON PIU' ESEMPI NEI NEIGHBOR

- 1NN CLASSIFIER o  $K=1$

- VORONOI TESSELTION

- CONDENSED 1NN

- USO UN SUBSET CHE NON ADICE PERFORMANCE
- CERCO UN CONSISTENT SUBSET E SCARTO RESTO
- POSSO REGOLARIZZARE CON GRANDEZZA SUBSET

- NON PARAMETRIC NEAREST MEAN

- DISTANZA EUCLIDEA o MAHALANOBIS

- NON PARAMETRIC MIXTURE OF GAUSSIANS

### DISTANCE LEARNING

- DIVERSE MISURE DI DISTANZA PER NEIGHBORHOOD
- LOCALLY ADAPTING DISTANCE FUNCTIONS
- PARAMETRIZZO DISTANZA E APPRENDO

- LARGE MARGIN NEAREST NEIGHBOR

- APPRENDO M di MAHALANOBIS

- SIMILARITY BASED REPRESENTATION

### OUTLIER DETECTION

- LOCAL OUTLIER FACTOR

- CONFRONTO DISTANZA K DI UN ELEMENTO CON SOMMA DELLE DISTANZE K DI TUTTI ELEMENTI NEL NEIGHBOR

### REGRESSION (SMOOTHER)

- MEDIA ELEMENTI NEIGHBOR PER PREVISIONE

- REGRESSOGRAM

- COME HISTOGRAM MA CON MEDIA  $y$

- RUNNING MEAN SMOOTHER

- COME NAIVE MA SIMMETRICO SU X

- KERNEL SMOOTHER

- KNN-SMOOTHER

- RUNNING LINE SMOOTHER

- NEIGHBORS PER REGRESSION LINE
- E USO PESO KERNEL INVECE DI NEIGHBORS

### PARAM OPTIMIZATION

- SMOOTHING SPINES

## 9 - DECISION TREE

- COSA È UN DECISION TREE
  - DECISION NODE
  - LEAF NODES

### ■ ALBERI UNIVARIATI

- USO UNA SOLA FEATURE
- DEVO DISCRETIZZARE FEATURE NUMERICHE  $x_i > w_{m0}$ 
  - ↳ BINARY SPLIT
- GREEDY CERCO BEST SPLIT

#### • CLASSIFICATION TREE

- VALUTO SPLIT CON IMPURITÀ
  - ↳ SE PURO NON SPLITTO ULTERIORMENTE

ENTROPY  
GINI INDEX  
MISCLASSIFICATION ERROR

↳ PROPRIETÀ IMPURITÀ

- IMPLEMENTAZIONI ID3 o C4.5
- PENALIZZARE ATTRIBUTI MOLTE FEATURE
- SOGLIA PUREZZA

#### • REGRESSION TREE

- UTILISMO MSE o Worst Possible Error
- PIECEWISE CONSTANT APPROXIMATION
- ERROR TREASURY
- NELLE FOGLIE POSSO AVERE LINEAR REGRESSION

#### • PRUNING

PREPRUNING - EARLY STOP  
POSTPRUNING  
↳ PRUNING SET

### ■ RULES

- RULE EXTRACTION
  - FEATURE EXTRACTION / INTERPRETABILITY
  - ↳ RULE BASE
    - RULE SUPPORT
    - RULE PRUNING

#### • RULE LEARNING

- DEPTH FIRST SEARCH
- UNA REGOLA ALLA VOLTA
- SEQUENTIAL COVERING - RIPPER
  - 1) ↳ MASSIMIZZO GAIN DI FOGL
  - 2) AGGIUNGO CONDIZIONI FINCHE NON CORRE NEGATIVI
  - 3) PRUNO MASSIMIZZANDO RULE VALUE METRIC
  - 4) STOPPING USANDO MDL
  - 5) PRUNING RULE BASE
  - 6) RULE OPTIMIZATION
- PER  $k > 2$  PASSO SEQ 2-CASE PROB

#### • FIRST ORDER RULES

- POSSONO CONTENERE PREDICATI
- DEFINISCO RELAZIONI ATTRIBUTI
- APPENDO TRAMITE LOGIC PROGRAMMING
- BINDING

### ■ MULTIVARIATE

- NODO INTERNO COME FIT LINEARE CON CONDIZIONE
  - ↳ LA UNIVARIATE NODE COME
- RICERCA ESAUSTIVA NON POSSIBILE
- POSSO USARE MONO NON LINEARE

↳ QUADRATICO  
↳ MLP  
↳ SPECKO

ALGORITHMI / CART  
OG1  
ASSUNDO CLASSI GAUSSIANE  
⋮

# 10 - LINEAR DISCRIMINATION

- APPROCCIO DISCRIMINANT BASED
  - └ ASSUNZIONI
  - └ BOUNDARIES
- LINEAR DISCRIMINANT
  - └ QUADRATIC
  - └ HIGHER ORDER TERMS
  - └ BASIS FUNCTIONS
- INTERPRETAZIONE GEOMETRICA
  - └ VETTORE PESI  $z$   $w_0$  COME SOGLIA
  - └  $x$  COME DISTANZA PROIEZIONE
  - └ ASSUNZIONE SEPARAZIONE LINEARE
- LINEAR CLASSIFIER
  - └ PAIRWISE SEPARATION
- PARAMETRIC DISCRIMINATION
  - └ LOGIT TRANSFORMATION - LINEARITA'
  - └ FUNZIONE LOGISTICA (SIGMOIDE)
- GRADIENT DESCENT
- LOGISTIC DISCRIMINATION
  - └ ASSUNZIONI
  - └ CROSS ENTROPY
  - └ GD
  - └ 2-NORMALIZATION
  - └ WEIGHT INITIALIZATION
  - └ EARLY STOPPING
  - └ SOFTMAX - PROPRIETA'
  - └ VARIANTI CON QUADRATIC - CONFRONTO PARAMETRIC
- MULTILABEL LOGISTIC DISCRIMINATION
  - └ SEPARATE 2-CLASS
- RANK LEARNING
  - └ SCORE FUNCTION
  - └ ERRORE



# 11 - MULTILAYER PERCEPTRON

- **PERCEPTRON**
  - WEIGHT & BIAS
  - TRESHOLD FUNCTION
  - LAYER
  - DIMENSIONALITY REDUCTION
  - ONLINE LEARNING - SGD
- **MULTI LAYER PERCEPTRON**
  - BACK PROPAGATION
    - MINIBATCH
    - BATCH LEARNING
    - CHAIN RULE
    - ERON
- **UNIVERSAL APPROXIMATION THEOREM** (MULTILAYER COMMON BIAS)
  - APPROXIO LOGICO
  - ESTENSIONE A R
  - PIECEWISE CONSTANT APPROXIMATION
- **OVERTRAINING - EARLY STOPPING**
- **HIDDEN REPRESENTATION LEARNING**
  - EMBEDDING
  - TRANSFER LEARNING
  - SEMI SUPERVISED LEARNING
- **AUTOENCODERS**
  - RECONSTRUCTION ERROR
  - NON LINEAR DIM REDUCTION
  - DENOISING
  - SPARSE
  - SAMMON MAPPING LEARNING

# 13 - LOCAL MODELS

- COMPETITIVE LEARNING
  - WINNER TAKE ALL
  - VANTAGGI ONLINE
- ONLINE K-MEANS
  - SDG
  - STABILITY PLASTICITY DILEMMA
  - RUN IMPLEMENTATION
    - LATERAL INHIBITION
      - EXCITATORY
      - INHIBITORY
  - HEBBIAN LEARNING - VANT / SVANTI
  - EMERGE DEAD CENTERS
- ADAPTIVE RESONANCE THEORY
  - VIGILANCE
- SELF ORGANIZING MAPS
  - CENTERS NEIGHBOR  $\rightarrow$  FUNCTIONE NEIGHBOR
  - TOPOGRAPHICAL MAP
- RADIAL BASIS FUNCTION
  - RAPPRESENTAZIONE LOCALE
  - RECEPTIVE FIELD
  - CALCOLO  $m_h$  e  $s_h$
  - GAUSSIAN UNITS + MLP
  - HYBRID LEARNING  $\sigma$  ANCHOR
    - DETERMINO  $m_h / s_h$
    - CALCOLO  $ph_e$
    - ADDESTRIO MLP CON INPUT  $ph_e$
  - DEFAULT MODEL con  $w_0$
- INCORPORARE RULE KNOWLEDGE
  - PRIORI
  - REGOLE DENTRO UNITA RBP
  - FUZZY MEMBERSHIP FUNCTION
- NORMALIZED BASIS FUNCTION
  - LINGUA DIPENDENZA
- COMPETITIVE BASIS FUNCTION
  - OUTPUT E MIXTURE MODEL
  - EM SUPERVISIONATO
- LEARN VECTOR QUANTIZATION
  - ASSEGNAMENTO e ALLONTANAMENTI
- MIXTURE OF EXPERTS
  - COOPERATIVE
  - COMPETITIVE
  - HIERARCHICAL MOE
    - SOFT DECISION TREE
  - GATING
    - EXPERTS

## 14 - KERNEL MACHINES

- SUPPORT VECTOR MACHINES
  - VANTAGGI
    - SUPPORT VECTORS
    - KERNEL FUNCTIONS
- IPERPIANO OTTIMALE SEPARAZIONE
  - DATA DA QUESTO LATO
  - ABBASTANZA LONTANI  $\rightarrow$  MARGINE
  - APPLICO MOLTIPLICATORI LAGRANGE
  - USO KKT
  - MASSIMIZZO SU MOLTIPLICATORI
    - $\rightarrow$  INTERPRETAZIONE
- SOFT MARGIN HYPERPLANE
  - VOGLIO IPERPIANO MINOR ERRORE
  - VARIABLE SLACK DEVIAZIONE MARGINE
    - LA SOFT ERROR  $\rightarrow$  PENALTA
  - LAGRANGIANA
    - INTERPRETAZIONE SUPPORT VECTOR
    - HINGE LOSS
- V-SVM
  - USO  $V$  invece di  $C$
  - SCALING MARGINE
  - CONTROLLA NUMERO SUPPORT VECTOR
- KERNEL TRICK
  - CONSTRAINT IN NUOVO SPAZIO
  - USO BASIS FUNCTION
  - RINPIAZZO INNER PRODUCT BASIS CON KERNEL
    - $\rightarrow$  GRAM MATRIX

### VECTORIAL KERNELS

- POLYNOMIALS  $\rightarrow$  LINEAR KERNEL
- RBF
- MAHALANOBIS
- GENERALIZED DISTANCE
- SIGMOIDAL

### CUSTOM KERNELS

- BAG OF WORDS
- EDIT DISTANCE  $\rightarrow$  EMPIRICAL KERNEL MAP
- DIFFUSION KERNEL
- FISHER KERNEL

### KERNEL COMBINATION

- COMBINAZIONE
- PESATA
- LOCALIZED KERNEL

### MULTICLASS KERNEL MACHINE

- K 2-CLASS
- SIGMOID OUTPUT
- SOFTMAX TRAINING
- PAIRWISE CLASSIFIERS
- SINGLE MULTICLASS

### KERNEL REGRESSION

- E-SENSITIVE LOSS
  - SLACK PER POSITIVE E NEGATIVE DEVIATION
  - AL CONTRARIO NEL TUBO  $\rightarrow$  GOOD FIT

### KERNEL RANKING

- DIFFERENZA RANKING COME CONSTRAINT
- DEFINITO COME 2-CLASSIFICATION

### KERNEL DENSITY ESTIMATION

- BOUNDARY SEPARA ZONE ALTA DENSITA'
  - $\rightarrow$  OUTLIER DETECTION
- VOGLIO INCLUDERE MAGGIORE DENSITA' IN SFERA
- $\alpha = 0$   $\rightarrow$  Istanza DENTRO
- KERNEL VA OLTRE SFERA

### LARGE MARGIN NN CLASSIFIER

- APPENDIMENTO MISURA DISTANZA
- KNN COME RANKING
- MARGINE DISTANZA NEIGHBOR CLASSI DIVERSE
- DISTANZA MAHALANOBIS
  - $\rightarrow$  OTTIMIZZO COVARIANZA
  - $\rightarrow$  REGOCARIZZO CON FATTORIZZAZIONE
  - $\downarrow$
  - LMCA

### KERNEL DIM REDUCTION

# 15 - GRAPHICAL MODELS

- MODELLI GRAFICI
  - ↳ JANTAGGI
  - ↳ DAG
  - ↳ CAUSAL GRAPH
- INCLUSIONE HIDDEN VARIABLE
- CASI CANONICI
  - ↳ HEAD TO TAIL → BLOCKING
  - ↳ TAIL TO TAIL
  - ↳ HEAD TO HEAD
    - ↳ EXPLAINING AWAY
- CASO GENERARE
- MODELLO GENERATIVO
  - ↳ CLASSIFICAZIONE
  - ↳ CLUSTERING
  - ↳ REGRESSIONE LINEARE
- d-SEPARATION
- MARKOV RANDOM FIELD
  - ↳ CLIQUE → MAXIMAL
  - ↳ POTENTIAL FUNCTION
  - ↳ MORALIZATION
- FACTOR GRAPH
  - ↳ SUM PRODUCT ALGORITHM
  - ↳ MAX PRODUCT ALGORITHM
- STRUCTURE LEARNING - PARAMETERS
  - ↳ STRUCTURE



# 16 - HIDDEN MARKOV MODELS

- DISCRETE MARKOV PROCESSES - PRIMO ORDINE

- TRANSIZIONE
  - AUTOMA STOCASTICO
  - PROBABILITA' INIZIALE

- OSSERVABILE  $\rightarrow$  APPRENDIMENTO PARAMS

- HIDDEN MARKOV MODEL
  - L PROBABILITA' OSSERVAZIONE

## ① VALUTAZIONE PROB OSSERV.

- PROCEDURA FORWARD - BACKWARD

- VAR  $\alpha$
  - VAR  $\beta$

(RIDEFINITO SENZA SEQ STATI)  
(MI BASTA  $\alpha$  PER QUESTO PROB)

## ② TROVARE SEQUENZA STATI

- RIDEFINISCO IN TERMINI DI  $\alpha$ ,  $\beta$

- VITERBI ALGORITHM  $\rightarrow$  DYNAMIC PROGRAM

## ③ APPRENDIMENTO PARAMETRI

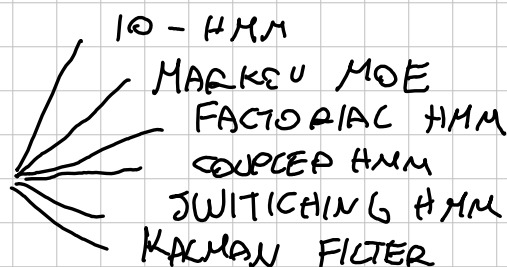
- USO  $\Sigma$  E DEFINITO IN TERMINI DI  $\alpha$ ,  $\beta$ 
  - $\hookrightarrow$  SOFT COUNT

- BAUM WELCH ALGORITHM

- E Step
  - M Step

- OSSERVAZIONI CONTINUE  $\rightarrow$  DISCRETIZZARE

- HMM COME MODELLI GRAFICI



- MODEL SELECTION
  - LEFT TO RIGHT HMM
  - HYBRID LR-HMM

# 17 - BAYESIAN ESTIMATION

- USO TUTTE LE PRIORI
  - └ M2
  - └ MAP
  - └ VANTAGGI
- └ A FULL BAYESIAN TREATMENT
  - └ CONJUGATE PRIOR → VANTAGGI
  - └ INTERPRETAZIONE PARAM.

## □ DISCRETE

- $K > 2$  → DIRICHLET
- $K = 2$  → BETA

## □ GAUSSIAN

### — UNIVARIATE

- + MEAN - VARIANCE → NORMAL
- - MEAN - VARIANCE → NORMAL-GAMMA

### — MULTIVARIATE

- - MEAN - VARIANCE → NORMAL-WISHART

## □ PARAMETERS FUNCTION

- REGRESSION → NORMAL (NOISE)
- REGRESSION (PRIOR NOISE PREC) → NORMAL-GAMMA

## □ BAYESIAN CLASSIFICATION

- $\pi^e | x^e \sim \text{BERNOULLI}$

### └ LAPLACE APPROXIMATION

- SCELTA PRIORI
  - └ LIVELLO 1
  - └ LIVELLO 2 (PARAMETRIZZATA)
  - └ LIVELLO 3 (VALUTATO SU DATI)

- INCLUDE BIAS / KERNEL

- MODEL COMPARISON
  - └ MODEL EVIDENCE
  - └ BAYESIAN FACTOR
  - └ BIC
  - └ AIC

## □ MIXTURE MODEL ESTIMATION

- MIXTURE PROPORTION → DIRICHLET
- MIXTURE COMPONENT → NORMAL-WISHART
- VARIATIONAL APPROXIMATION

## □ MODELLO A ZIONE NON PARAMETRICO

- PROCESSI
  - └ GAUSSIAN
  - └ DIRICHLET - LATENT DIRICHLET ALLOC
  - └ BETA

# 19 - ENSEMBLES

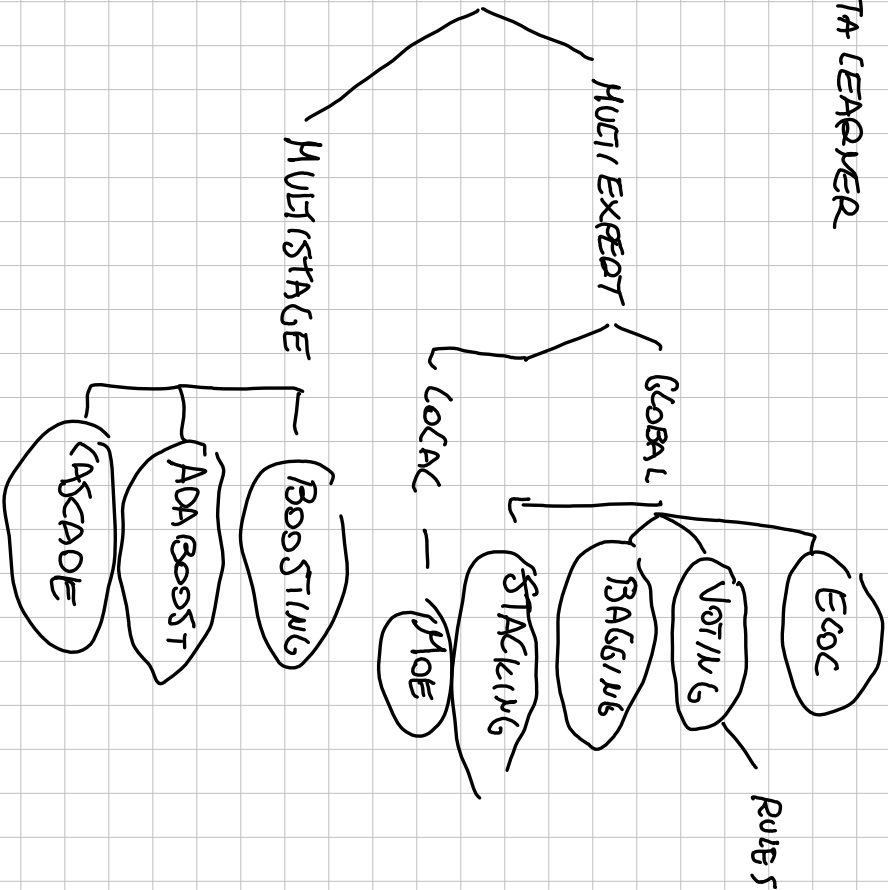
- DIVERSE LEARNERS

↳ DIVERSITY & ACCURACY

DIFFERENT {  
ALGORITHM  
HYPERPARAMETERS  
INPUT REPRESENTATION  
TRAINING SET

- FINE TUNE { ENSEMBLE SELECTION  
META LEARNER

- MODE COMBINATION



20 - EVAGUATION