

Indice

PREMESSA	Pag.	1
1. Informazione ed Incertezza in Statistica	“	1
2. Un esempio: Sperimentazione Clinica	“	4
3. Misure di Incertezza	“	7
 PARTE 1: ANALISI DELLA DIPENDENZA		
Capitolo 1: ASPETTI GENERALI	“	17
1. Premessa	“	17
2. Notazioni	“	17
3. Problema, Informazione ed Incertezza	“	17
3.1 Informazione Empirica (Dati)	“	17
3.2 Obiettivo	“	18
3.3 Incertezza	“	18
4. Approcci metodologici (quadro generale)	“	18
4.1 Minimi Quadrati	“	18
4.2 K-Nearest Neighbour	“	20
4.3 Valutazione dell'Informazione e dell'Incertezza	“	21
 Capitolo 2: METODI LINEARI DI ANALISI DELLA DIPENDENZA	“	25
A) MODELLO LINEARE CLASSICO	“	25
1. Struttura del modello	“	25
2. Stima MQ	“	26
3. Analisi dell'Incertezza	“	27
 B) MANIPOLAZIONI DEL MODELLO LINEARE CLASSICO	“	34
4. Motivazioni	“	34
5. Selezione di sottoinsiemi di Predittori	“	34
5.1 Forward Stepwise Selection	“	35
5.2 Backward Stepwise Selection	“	35
6. Contrazione dei Coefficienti	“	35
6.1 Regressione Ridge	“	36
6.2 Metodo “Lasso”	“	38

Capitolo 3: METODI NON LINEARI DI ANALISI DELLA DIPENDENZA	Pag.	41
1. Aspetti generali	“	41
A) Tecniche basate sulla RSS penalizzata	“	42
B) Metodi del Nucleo e della Regressione Locale	“	42
C) Tecniche fondate sulle Funzioni Basali e Metodi del Dizionario	“	43
2. Metodi del Nucleo	“	45
2.1 Caso $m=1$	“	45
2.2 Caso $m>1$	“	50

PARTE 2: CLASSIFICAZIONE

Capitolo 4: METODI DI CLASSIFICAZIONE:	“	57
1. Aspetti generali	“	57
2. La Classificazione dal punto di vista statistico	“	58
3. I due problemi di Classificazione	“	58
4. Richiami sulla Classificazione Non Supervisionata	“	59
4.1 La rappresentazione matematica dei dati	“	59
4.2 Tipi di tecniche di Cluster Analysis	“	60
 Capitolo 5: CLASSIFICAZIONE SUPERVISIONATA	“	69
1. Introduzione	“	69
A) METODI LINEARI DI CLASSIFICAZIONE	“	73
2. Modello di Regressione Lineare per una Matrice di Variabili Indicatrici	“	73
3. Analisi Discriminante Lineare	“	76
3.1 Analisi Discriminante Quadratica	“	78
3.2 Analisi Discriminante Regolarizzata	“	78
3.3 Analisi Discriminante Lineare a Rango Ridotto	“	79
3.4 Approccio Fisheriano	“	81
4. Modello Logistico di Regressione	“	83
5. Metodi basati sulla Modellizzazione Lineare delle Frontiere di Classificazione	“	90
5.1 Premessa	“	90
5.2 Algoritmo del Percettrone di Rosenblatt	“	92
5.3 Iperpiani Ottimali di Separazione	“	94

B) METODI NON LINEARI DI CLASSIFICAZIONE	Pag.	100
6. Metodi del Nucleo	“	101
6.1 Stima delle Funzioni di Probabilità nei Gruppi	“	101
6.2 Regole di Classificazione basate sul Metodo del Nucleo	“	102
7. Analisi Discriminante con Modelli Mistura	“	105
8. Metodo della “Support Vector Machine”	“	112
8.1 Dall’approccio lineare al non lineare	“	112
8.2 L’approccio non lineare	“	116
9. Alberi di Classificazione	“	120
9.1 Alberi di Decisione	“	120
9.2 Alberi di Classificazione	“	122
9.3 Costruzione di un albero di classificazione (“ottimale”)	“	124
9.4 Valutazione della performance di un albero deterministico di classificazione	“	131
9.5 Stime di $R(T)$	“	132
9.6 Costruzione dell’albero “ottimale” in base alla potatura	“	134
9.7 Misura d’importanza di una variabile	“	136
10. Reti Neurali	“	137
10.1 Nozioni introduttive	“	138
10.2 Formalizzazione matematica di una rete neurale	“	139
10.3 Processo di stima di una RN	“	143
10.4 RN e Classificazione	“	148

PARTE 3: GESTIONE DELL’INCERTEZZA

Capitolo 6: METODI DI GESTIONE DELL’INCERTEZZA NELLA REGRESSIONE E NELLA CLASSIFICAZIONE SUPERVISIONATA	“	153
1. Introduzione	“	153
2. Decomposizione degli Errori	“	158
2.1 Bias e Varianza	“	158
2.2 Ottimismo dell’Errore di Training	“	160
3. Metodi basati sulla stima di Err_{in}	“	163
3.1 Statistica C_p	“	163
3.2 Akaike Information Criterion	“	163
3.3 Bayesian Information Criterion	“	166
3.4 Dimensione di Vapnik-Chervonenkis	“	166
4. Metodi basati sulla stima diretta di Err	“	170
4.1 Validazione incrociata	“	170
4.2 Metodi Bootstrap	“	172

PARTE 4: ANALISI DELLA STRUTTURA LATENTE

Capitolo 7: MODELLI DI ANALISI DELLA STRUTTURA LATENTE	Pag.	179
1. Introduzione	“	179
2. Modelli delle classi latenti	“	182
2.1 Stime di massima verosimiglianza	“	184
2.2 Attendibilità delle stime di massima verosimiglianza	“	188
2.3 Test di bontà dell’adattamento del modello	“	189
2.4 Assegnazione degli individui alle classi latenti	“	190
2.5 Interpretazione delle classi latenti	“	190
3. Modelli dei profili latenti	“	196
4. Modelli fattoriali lineari	“	197
4.1 Stime di massima verosimiglianza	“	199
4.2 Attendibilità delle stime di massima verosimiglianza	“	200
4.3 Test di bontà dell’adattamento del modello	“	200
4.4 Valutazione degli individui in termini fattoriali	“	201
4.5 Interpretazione dei fattori	“	202
4.6 Un esempio	“	204
Capitolo 8: MODELLI AD EQUAZIONI STRUTTURALI	“	209
1. Introduzione	“	209
2. I Gruppi di Variabili	“	210
3. Modello SEM	“	212
4. Assunzioni sul modello SEM	“	214
5. Parametri del modello SEM	“	216
6. Relazione fondamentale dei modelli SEM	“	216
7. Stima dei parametri del modello SEM	“	217
8. Un esempio: Health Attainment	“	218
BIBLIOGRAFIA	“	227