

oggetto, ne elabora le *features* per identificare il *pattern*, e restituisce un valore discreto che determina a quale categoria appartiene.

In generale gli approcci si suddividono in generativi e discriminativi.

Il generativo mira a definire un modello per ogni categoria, o classe, da discriminare nel dataset. Questa forma presenta una struttura più flessibile in grado di adattarsi a nuove classi, è più rapida nell'addestramento e ottiene una migliore capacità descrittiva per la singola classe.

L'approccio discriminativo, invece, si basa sulla ricerca del migliore confine decisionale per separare le classi nello spazio. Per sua natura è più rapido nella scelta quindi velocizzando quindi la fase di test. Risulta molto efficace, ed è flessibile nell'eventualità di un modello errato. cosa vuol dire?

Un'ulteriore suddivisione nei classificatori si basa sulla loro natura parametrica o non parametrica. La parametrica si caratterizza per l'assunzione della forma di distribuzione dei dati per ogni classe (es. la distribuzione normale), e si concentra nella stima dei parametri della funzione che genera tale distribuzione. Diversamente, la non parametrica non assume nessuna forma di distribuzione, ma viene stimata direttamente dal *training set*. Risulta più dispendiosa in termini computazionali ma non basandosi su assunzioni determina un modello maggiormente flessibile e adattabile al contesto.

2.4.1 Apprendimento supervisionato

Nella PR spesso si utilizza il paradigma dell'*apprendimento da esempi*. Questo metodo può essere visto come l'apprendimento di un bambino che sperimenta e acquisisce conoscenza da esempi, dall'esperienza. In particolare, nel contesto della classificazione, si utilizza un approccio *supervisionato* in cui la conoscenza viene acquisita tramite dati conosciuti, il *training set*, dotato di categorie, o etichette, note. Conoscendo la reale classe di appartenenza degli oggetti, il modello può addestrarsi e affinare gradualmente la sua capacità decisionale.

In questo processo è importante che il sistema non "impari a memoria" i dati, il cosiddetto *overfitting*, un eccessivo adattamento ai dati di addestramento. Per evitare ciò il modello deve essere in grado di generalizzare quindi di classificare correttamente anche oggetti sconosciuti, mai visti. In tal modo dimostra di non essersi specializzato sui dati di addestramento e di aver appreso le caratteristiche essenziali.

L'obiettivo è
infatti di
creare un

2.3.2 Validazione

predetto dal classificatore

Una volta costruito il modello è necessario verificare la qualità del classificatore. Per tale scopo si utilizza il *testing set*, un *dataset* che presenta elementi sconosciuti al sistema, quindi diversi ~~per~~ da quelli usati in addestramento, ma dotato di categorie note da confrontare con il risultato. Viene eseguito il modello su tali dati e si valuta la predizione in base all'errore ottenuto. Questo valore pone in rapporto le previsioni errate rispetto al numero totale di oggetti analizzati. Una previsione errata consiste in una falsa valutazione del classificatore, quindi un valore diverso dalla reale categoria di appartenenza.

Nella costruzione di un classificatore di solito si dispone di un unico *dataset* che si deve suddividere in due parti, il *training set*, per l'addestramento, e il *testing set* per i test e la validazione. Un metodo ideale sarebbe poter usare tutti dati di esempio per il *training* ed estrarre altri esempi dal problema per testare il modello, ma nella realtà potrebbe essere non fattibile o troppo dispendioso. Si preferisce quindi suddividere il *dataset* in due parti uguali, una proporzione 50/50, nella modalità definita come "*train-test split*". Tuttavia, una *metodica comune* consiste nella *cross validation*, che permette di ottenere una valutazione più valida e consistente. Esistono diverse varianti, ognuna con le sue caratteristiche. La forma più semplice è la *holdout*, che distribuisce casualmente i dati in due insiemi di uguale dimensione. Un'alternativa simile è l'*average holdout*, che per essere indipendente dalle partizioni effettua più *holdout* e calcola l'errore come media dei risultati ottenuti in tutti i casi.

Infine, una delle più utilizzate è la *Leave One Out* (LOO), una variante particolare che ottiene ottimi risultati in termini di affidabilità, soprattutto con *dataset* ristretti. Come suggerisce il nome, consiste nell'effettuare l'addestramento con tutti gli oggetti del *dataset* meno uno, x_i , che viene invece utilizzato per validare il modello. Si ripete il procedimento per ogni elemento x_i del *dataset*, e al termine si media il risultato ottenuto. Presenta un costo computazionale maggiore ma garantisce indipendenza dalla partizione e dai dati scelti del *training set* e *testing set*.

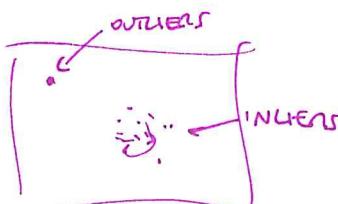
Leave one out

Leave one out
average holdout
outlier
inlier

2.5 Anomaly Detection

L'*anomaly detection* (AD) consiste nell'identificare fenomeni ed eventi che presentano un comportamento anomalo rispetto al resto dei *dataset*. Tali fenomeni, denominati *outlier*, si confondono tra gli *inlier*, il resto dei dati normali, per la loro caratteristica natura e la scarsa numerosità.

majority metta 1 esempio tip



2.5.1 Gli outlier

Individuare gli *outlier*, pur essendo complicato, è una necessità. Ciò consentirebbe di inferire informazioni preziose oppure rappresentare una prevenzione per eventuali criticità. Contribuisce in modo significativo anche nel *data cleaning*, la "pulizia dei dati", essenziale per la preparazione dei dati per l'analisi e altre svariate applicazioni. Si tratta di un insieme di processi che servono a rimuovere duplicati, uniformare e filtrare i dati. In questo contesto rimuovere gli *outlier* semplificherebbe le fasi successive di analisi migliorando la qualità del risultato.

Si distinguono varie tipologie di *outlier*:

- i *point*, che sono identificati tali indipendentemente se si trovano da soli o in gruppo;
- i *collective*, considerati solo se rilevati in gruppo, altrimenti rientrano negli *inlier*;
- i *contextual*, dove il contesto determina la loro forma, normale o anomala; *non
diametrali*

Mogli' fai un esempio

2.5.2 Le applicazioni

L'AD ricopre un ruolo importante in molteplici campi. Per esempio, possiamo citare:

- le intrusioni di rete, dove l'attaccante deve mantenere un basso profilo per non essere individuato all'interno del sistema; *non
belkr.
simi*
- l'ambito sanitario, dove rilevare un segnale anomalo nelle diagnosi potrebbe indicare malattie gravi, come il cancro;
- i sistemi automatizzati, in cui prevenire un avaria o, riuscire a intervenire in tempo nel caso si manifestasse, è essenziale;
- nel processamento di immagini o testi, come rilevare *fake news*;
- nel rilevamento di frodi, all'interno delle innumerevoli transazioni bancarie prodotte ogni giorno.

Queste casistiche sono accomunate da un'enorme quantità di dati dove sistemi troppo rigidi e specifici non potrebbero adattarsi al continuo mutamento delle variabili in gioco. Le problematiche nell'AD consistono non solo in dataset di grandi dimensioni, ma anche dall'alto costo computazionale derivato per elaborarli. Inoltre, un'eccessiva numerosità di *features* potrebbe portare alla *curse of dimensionality*, un fenomeno dove un'alta dimensionalità comporta una drastica riduzione delle prestazioni.

*è la
stessa
cosa?*

questo non è specifico x AD,

ma vale x ogni classificazione o riconoscimento di PR

Se vuoi metti qualcosa d'specifico x AD 13

(non superatevi),

non ha senso d'averli, etc etc)

2.5.3 I metodi

I diversi approcci di AD tipicamente suddividono nello stesso modo
In generale la suddivisione si basa sul metodo utilizzato per rilevare anomalie. I vari metodi
possono essere:

- Metodi basati sulla statistica: essi ricercano elementi che non rispecchiano la distribuzione dei dati. Una bassa probabilità di appartenenza alla forma di distruzione determina un'alta probabilità di essere un outlier.
- Metodi basati sul clustering: essi suddividono i dati per somiglianza in gruppi, i *cluster*, e le anomalie risultano evidenti poiché troppo diverse dal loro gruppo di appartenenza. Oppure gli outlier definiscono gruppi malformati con altre anomalie.
- Metodi basati sull'apprendimento: vengono utilizzati algoritmi di apprendimento automatico specializzati a modellare il comportamento normale e individuare le anomalie. (non dicono niente)
- Metodi basati sulla distanza o sulla densità: viene considerata la distanza tra gli elementi, o la densità locale. Nel primo caso gli *outlier* si troveranno distanti dagli altri punti, nel secondo saranno in zone a bassa densità.
- Metodi basati su insiemi, o ensemble: queste tipologie combinano i risultati di metodi diversi, o gli stessi con parametri differenti, per ottenere una previsione più accurata.

La maggioranza degli algoritmi rappresenta il proprio risultato mediante un valore, denominato *anomaly score*, che quantifica quanto un elemento è probabile che sia un'anomalia.

In letteratura esistono differenti algoritmi di AD, ma di seguito ne saranno descritti solo tre, quelli utilizzati nello studio. I metodi sono L'IForest (IF), o Isolation forest, il Local Outlier Factor (LOF) e l'Ocsvm (OCSVM), o One Class Support Vector Machine.

IF è un algoritmo basato su ensemble, e si sviluppa su alberi decisionali. Per albero decisionale si intende una struttura ad albero dove ogni nodo rappresenta una decisione e i rami le possibili alternative. Procedendo dall'alto verso il basso del modello, i dati vengono suddivisi nelle varie decisioni fino alle foglie. In IF viene costruita una foresta di alberi decisionali attraverso una selezione casuale dei dati. Ciascun albero cerca di isolare i dati mediante suddivisione, e le anomalie si troveranno più in alto poiché avranno bisogno di meno divisioni rispetto ai dati normali. In sostanza, più risulta facile isolare un oggetto e con maggiore probabilità sarà un'anomalia. Questo algoritmo risulta molto scalabile, veloce e con una bassa tendenza all'overfitting.

può essere calcolata guardando

LOF è un metodo basato sulla densità. La densità di un oggetto dipende dal suo vicinato, ovvero dalla numerosità degli elementi che gli sono vicini. Il confine per definire vicino un oggetto dipende dalla metrica scelta nell'implementazione. La più comune è la distanza euclidea, che misura la lunghezza del segmento tracciato tra due punti. In sostanza, il funzionamento definisce che minore è la densità locale più alta è la probabilità che ~~in~~ un determinato oggetto potrebbe esserci un'anomalia. Questo metodo è efficace con dati ad alta dimensionalità e nel rilevamento di anomalie, ma al contrario, risulta molto dispendioso in termini computazionali.

sia, parlando di questo

Infine, OCSVM si basa sull'apprendimento. Si tratta di una variante delle *Support Vector Machine*, un approccio discriminativo applicato solitamente a problemi binari. In questa forma, a singola classe, l'algoritmo viene addestrato per definire un confine che racchiude al suo interno solo i dati normali. Così facendo i dati anomali, che risultano esterni alla distribuzione, emergono e sono quindi rilevabili. Spesso viene utilizzato il trucco del *kernel*, che consiste nel proiettare i dati in una dimensione superiore, dove può risultare più facile separare i dati. Questo metodo presenta un buon adattamento anche a dati non lineari, grazie alla funzione *kernel*. In svantaggio si presenta meno scalabile per dataset molto grandi.

Se è detto

QUESTA PARTE
E' MOLTO DIVERSO
DALLE PRECEDENTI
COME STIE,
"PRESA DA QUALCHE PARTE" ?

NOTA generale: non puoi aggiungere delle figure esplicative
(non solo più in sede nei capitoli precedenti)

Capitolo 3

Dataset

Lo studio condotto si è basato sui dati raccolti nell'articolo della rivista scientifica "Biota Colombiana" pubblicata dall'*Instituto de Investigación de Recursos Biológicos Alexander von Humboldt*, che ha sviluppato un analisi di monitoraggio acustico passivo nella *Riserva Naturale Los Yátaros*, nel dipartimento di *Boyacá* in *Colombia* [2]. Con il termine passivo si identifica una modalità di osservazione del paesaggio incentrata sulla registrazione di un particolare luogo e solo successivamente prevede un'analisi approfondita, diversamente da quella attiva, dove si osserva e si analizza il fenomeno in tempo reale.

La riserva è composta da querceti e foresta subandina in diversi stadi di rigenerazione naturale, e presenta una biodiversità acustica molto particolare. Il progetto mirava a profilare l'impronta acustica della riserva campionando suoni nello spettro udibile e negli ultrasuoni. Sono stati predisposti tre siti, denominati *YAT*, lungo il sentiero principale, distanti 150 m, con due sensori acustici *AudioMoth* ciascuno, per le due forme di suono desiderate, posti ad altezza diverse. Il periodo di campionamento si è svolto dall'1 marzo al 2 maggio 2020, registrando un minuto di audio ad intervalli di trenta minuti, quindi 48 giornalieri, e per un totale di 9055 audio nell'udibile e 3392 nell'ultrasuono. Nel contesto applicativo, sono stati considerati solo i dati nello spettro udibile, per permettere l'ascolto del contenuto.

origine della differenza?

3.1 Dataset prima fase

Nella prima fase dello studio è stato utilizzato il *dataset* completo (DATA1) che presenta l'insieme originale dei dati. Il gruppo si presenta con una suddivisione per i tre siti (YAT1, YAT2, YAT3) con quantità leggermente differenti. Gli audio sono 3018 per YAT, a parte il primo con 3019. Ogni sito presenta 1482 file per il mese di marzo (1483 solo per YAT1), 1440 per il mese di aprile e 96 per il mese di maggio.

Data l'ingente quantità di dati disponibile, non si è potuto analizzare il contenuto, ossia ascoltare l'intero insieme di registrazioni. In un primo sondaggio, analizzando diversi audio in momenti diversi della giornata e del mese, si è esplicitato che tutti e tre i luoghi risultano molto caratterizzati dal suono del fiume e della cascata vicina. Anche se tale suono fosse stato ad una distanza maggiore avrebbe sortito lo stesso effetto. Farina et alii sostengono che "la geofonia può essere rilevata anche a grandi distanze in base all'ampiezza della sorgente

sonora” [2]. A questo si aggiungono ulteriori problematiche dovute a periodi piovosi, il cui rumore sovrasta in diverse occasioni i suoni ambientali naturali. Entrambi gli elementi appena descritti determinano un ambiente umido che potrebbe influire anche sulla capacità del sensore.

3.2 Dataset seconda e terza fase

Il *dataset* della seconda e terza fase dello studio (DATA2) consiste in un sottoinsieme del ~~dataset DATA 1~~ primo appena analizzato. Per potere inferire maggiori informazioni dal contesto si è stabilito che era necessaria una descrizione più accurata del contenuto. Quindi si è ristretto l'insieme ad un campione di dati minore che potesse essere ascoltato e studiato nel dettaglio. L'analisi ha ~~ha effettuato~~ ridefinito il *dataset* originale ~~in~~ 186 audio, focalizzandosi sul sito YAT1, nel mese di marzo con finestre temporali a intervalli di due ore, partendo dalle due del mattino, quindi nelle ore: 02:00, 06.00, 10:00, 14:00, 18:00, 22:00. Con questa modalità si poteva ottenere una visione abbastanza generale della varietà sonora presente nella giornata, includendo i due momenti fondamentali alba e tramonto, caratterizzati da picchi di attività acustica.

L'interpretazione manuale del *dataset* ha classificato il contenuto assegnando delle etichette ai vari elementi distinti. In relazione all'ANT è stato individuato un unico suono, appartenente al rumore di veicoli (classe V). Nella BIO è stato individuato il verso degli uccelli e dei grilli (classi U e G). Per la GEO si è rilevato il precedentemente menzionato rumore del fiume/cascata, la pioggia e i tuoni (classi C, P e T). Infine, sono stati identificati i rumori relativi alle interferenze del sensore (classe I), ed eventuali elementi uditi ma non interpretati, purtroppo sconosciuti (classe S). Rispetto a quanto specificato nell'introduzione, all'interno della GEO si è integrato anche l'insieme dei suoni relativi alla quiete. Tale scelta è derivata da una maggiore semplicità nella trattazione, ma specialmente per l'impossibilità nel poterli classificare correttamente.

Nel grafico 3.1 è possibile visualizzare la distribuzione degli elementi descritti nelle fasce analizzate. Da una prima osservazione risulta evidente la significativa presenza dell'elemento C, come già esplicitato nel paragrafo precedente. Lo accompagna il suono dell'elemento U, risultato meno attivo solo nella parte centrale della giornata. Il resto è relativamente distribuito, a parte I e S diffusi con bassa intensità. Si può notare come S sia presente solo nelle due fasce pomeridiane. Dal grafico 3.2, possiamo avere una visione del mese di marzo, per il quale è lecito esplicitare le medesime considerazioni annotate.

3 cosa vuol dire?

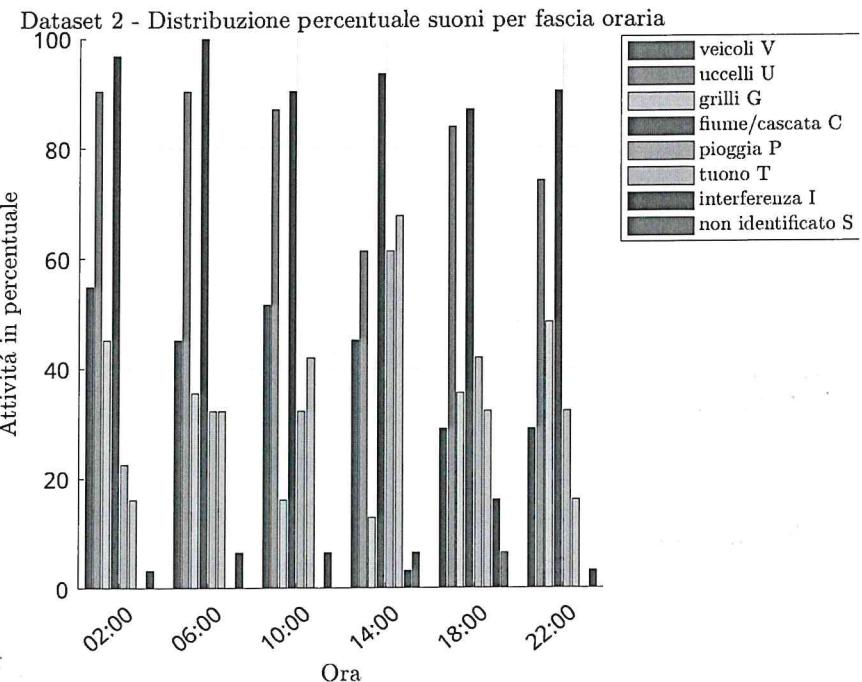


Grafico 3.1

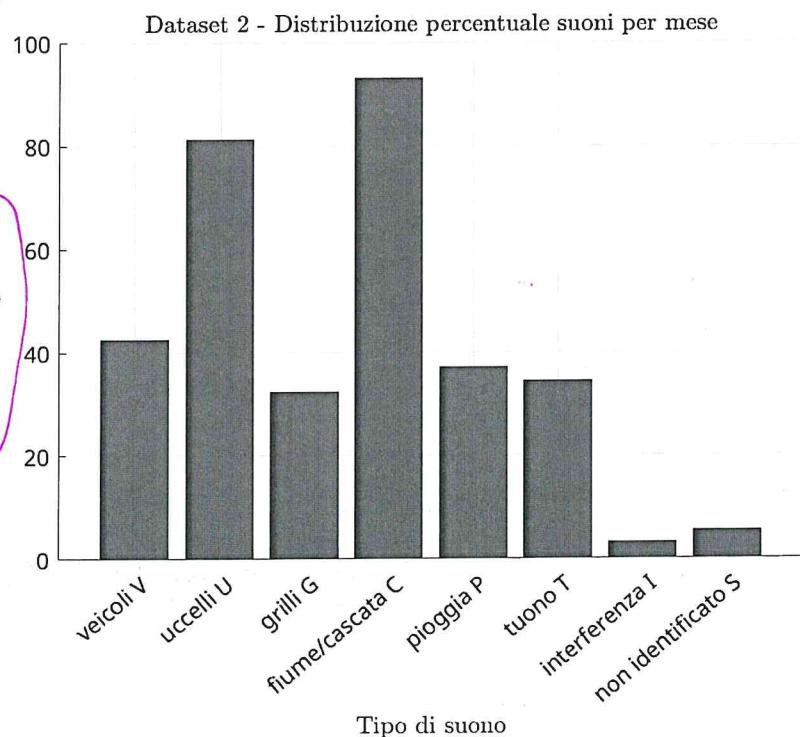


Grafico 3.2

Capitolo 4

→ fare x spiegare la differenza

Classificazione

veniamo decisi i risultati delle

In questo capitolo si desidera analizzare le prime due fasi dello studio basate sulla classificazione: la prima riguardante la rappresentazione su categorie note, la seconda esplorativa su categorie semantiche. Innanzitutto saranno definiti i gruppi di features utilizzati negli esperimenti, descrivendo poi la tipologia di classificatore scelto. A seguire, saranno illustrati nel dettaglio i problemi di classificazione disegnati, ed infine si andrà ad analizzare i risultati ottenuti nei due esperimenti.

→ per esplorare meglio x leggere

L'obiettivo è determinare quale finestra risulta più funzionale, quali features sono più efficaci al nostro contesto, e in quali problemi ottengono i risultati migliori. Questo studio considera gli esperimenti di classificazione precedentemente condotti dai colleghi Ilaria Ballerini e Andrea Piazza.

vive qualche
la possibilità d'agire

Inoltre, è stata valutata una fase di filtraggio per alcune frequenze, per capire se potesse rappresentare un miglioramento alla qualità dei dati in input. L'idea consiste nel rimuovere alcuni elementi rumorosi preponderanti su ogni audio, dovuti in maggior parte alle problematiche intrinseche del contesto, ma anche alla sensibilità dello strumento di misurazione.

4.1 Configurazioni features

Prima di poter costruire un modello di classificazione in grado di spiegare il contesto esaminato occorre identificare le caratteristiche necessarie che possano descrivere gli oggetti studiati. Nello studio sono state considerate sia le features nella loro forma originale, come descritta nel capitolo 2.2, sia raggruppate in insiemi. Gli elementi di gruppi sono stati derivati

attraverso concatenazioni delle caratteristiche originali o delle loro medie matematiche. Si premette che quelle identificate come originali in realtà sono descritte da un numero definito di componenti che per il momento verranno indicate con la variabile N , che ne indica la numerosità: maggiori dettagli sulla numerosità di N saranno forniti alla fine di questo paragrafo. Ogni feature originale è stata descritta da un vettore riga di $1 \times N$ componenti.

In particolare sono stati identificati 19 gruppi di features:

- 11 ORIGINALI (o ORIG), per ognuna delle 11 features, utilizzate singolarmente, $1 \times N$ elementi ciascuna;

- CONCATENAZIONE ORIGINALI (o CONC.ORIG.), formata dalla concatenazione orizzontale delle 11 *features* originali, ottenendo un totale di $11 \times N$ componenti;
- CONCATENAZIONE SPETTRALI (o CONC.SPE.), ottenuta dalla concatenazione delle componenti delle 5 *features* spettrali, quindi $5 \times N$ componenti;
- CONCATENAZIONE TONALI (o CONC.TON.), come la precedente ma considerando le 3 *features* della tonalità, ovvero $3 \times N$ componenti;
- CONCATENAZIONE TEMPORALI (o CONC.TEM.), come la precedente ma con le 3 *features* temporali, $3 \times N$ componenti;
- CONCATENAZIONE MEDIE ORIGINALI (o CONC.MED.ORIG.), formata dalla concatenazione orizzontale delle medie delle 11 *features* originali, quindi 11 componenti;
- CONCATENAZIONE MEDIE SPETTRALI (o CONC.MED.SPE.), ottenuta dalla concatenazione orizzontale delle medie delle componenti delle 5 *features* spettrali, quindi solo 5 componenti;
- CONCATENAZIONE MEDIE TONALI (o CONC.MED.TON.), come la precedente ma utilizzando le 3 *features* tonali, quindi solo 3 componenti;
- CONCATENAZIONE MEDIE TEMPORALI (o CONC.MED.TEM.), come la precedente ma utilizzando le 3 *features* temporali, in totale 3 componenti;

Ai 19 gruppi sopracitati si è tenuto conto anche della relativa versione standardizzata, ovvero ottenuta dalle componenti processate con la tecnica di standardizzazione *Z Score*, definendo quindi 38 gruppi: 19 *puri* e 19 standardizzati. In questo modo, si dispone anche di una rappresentazione con una scala comune indipendente dalle misurazioni originali.

In aggiunta, i dati sono stati estratti in due forme diverse (ottenendo quindi 78 gruppi di features) basate su diverse finestre temporali scelte per il campionamento nel calcolo dello spettrogramma: la prima si basa su un tempo minore di 1 secondo, che indicheremo con FS0X, a 32768 campioni; la seconda, FS1, invece si basa su un tempo uguale a 1 secondo, ovvero a 48000 campioni.

*[...] La precedente variabile definita come N è stata introdotta per indicare il numero di componenti per ogni *feature*. Per poterla definire era necessario introdurre il concetto relativo alle due finestre di intervallo misurate. Nel caso di FS1 sono state estratte 120 componenti ($N=120$): il segnale audio analizzato presenta una lunghezza temporale di 60 secondi e, per costruire lo spettrogramma lo si analizza con un intervallo di 1 secondo alla volta, ottenendo 60 finestre. Inoltre, considerando un passo pari alla metà dell'intervallo, 0.5 secondi, si*

Configurazione

ottengono altre 60 finestre, per un totale di 120 campionamenti. Alla stessa modo, è stato fatto per la forma FS0X dove, considerando una finestra più breve, si è ottenuto un maggior numero di componenti ($N=176$).

La tabella 4.1 riassume i vari gruppi di *features* considerati (il simbolo # indica la numerosità).

non lo mettere ..

GRUPPI DI FEATURES	DESCRIZIONE FEATURES	# COMPONENTI PER FEATURE	# N x FEATURE CASO FS0X	# N x FEATURE CASO FS1
CONC.ORIG.	Concatenazione originali	11	1320	1936
CONC.SPE.	Concatenazione spettrali	5	600	880
CON.TON.	Concatenazione tonali	3	360	528
CONC.TEM.	Concatenazione temporali	3	360	528
CONC.MED.ORIG.	Concatenazione medie originali	11	11	11
CONC.MED.SPE.	Concatenazione medie spettrali	5	5	5
CONC.MED.TON.	Concatenazione medie tonali	3	3	3
CONC.MED.TEM.	Concatenazione medie temporali	3	3	3
ORIG	Spectral Centroid	1	120	176
ORIG	Spectral Crest Factor	1	120	176
ORIG	Spectral Decrease	1	120	176
ORIG	Spectral Flatness	1	120	176
ORIG	Spectral Flux	1	120	176
ORIG	Spectral Roll off	1	120	176
ORIG	Spectral Spread	1	120	176
ORIG	Spectral Tonal Power Ratio	1	120	176
ORIG	Time Zero Crossing Rate	1	120	176
ORIG	Time Acf Coeff	1	120	176
ORIG	Time Max Acf	1	120	176

Tabella 4.1

4.2 Dettagli classificatore

In questo studio è stato utilizzato il classificatore *K Nearest Neighbor* (*KNN*), un approccio supervisionato generativo non parametrico, semplice e intuitivo: il metodo si basa sul classificare un punto, un oggetto, assegnandogli la classe che più frequentemente ritroviamo tra i k oggetti più vicini. Il concetto di vicinanza è relativo alla forma scelta di

*si concentra
con le
voci della
classe*

*sposta
del
capitolo
del
BG*

10 DIRE:

etichette note → classificazione?

"pensa
a qualcosa
di meglio")

specializzazione dell'algoritmo: nel nostro caso la distanza euclidea, una delle misure più utilizzate. Come risulta chiaro, la scelta del valore di k è cruciale. È stato utilizzato il valore più semplice, con k uguale a 1, facile da implementare e da comprendere, in grado di catturare dettagli molto fini nel dataset poiché nella decisione della classe si basa unicamente sull'elemento più vicino (comunemente viene indicata con solo 1-NN, o solo NN). A suo svantaggio, un valore troppo piccolo, come nel nostro caso, lo rende molto sensibile al rumore, determinando risultati errati o incongruenze, influenzando l'accuratezza del modello.

Per ovviare al problema, è stata integrata una validazione incrociata *Leave One Out* (LOO), per migliorare la robustezza e l'affidabilità del risultato. Per semplicità l'insieme dei due metodi sarà indicato con *LOO KNN*.

4.3 Problemi di classificazione disegnati

i due problemi

hanno caratteristiche

Le due fasi di classificazione di questo studio affrontano problemi diversi che illustreremo nei prossimi paragrafi, ma in sostanza si differenziano dalla tipologia di etichette scelte: la prima rappresenta i dati utilizzando delle etichette note, la seconda invece delle etichette semantiche.

cerca di separare dati

di natura

etichette note (pieno/nullo)

dati di natura

4.3.1 Rappresentazione con etichette note

Il dataset è

La prima fase propone uno studio sul dataset DATA1. Essendo privo di annotazione, procedere già dall'inizio con la classificazione supervisionata non era possibile; inoltre optare per un etichettatura manuale, identificando i vari suoni all'interno, non era considerabile, sia per l'eccessivo tempo necessario che per la mancanza di risorse. Tuttavia, come specificato nell'introduzione, i limiti più ostici consistono da una parte nella difficoltà oggettiva intrinseca di discriminare elementi all'interno di un *soundscape* e, dall'altra, nella competenza tecnica necessaria a identificare la biodiversità presente. Per questi motivi, si è deciso di proporre dei problemi affrontabili basati su etichette note, cioè su informazioni deducibili dal contesto dell'oggetto, invece che dal suo contenuto: si è tenuto conto del luogo di registrazione, e della temporalità, come l'ora del giorno, o una fase della giornata, o del mese.

Sono stati individuati i seguenti problemi:

- PR-1.1 YAT. Riguarda il luogo di registrazione, gli *yat*. Si tratta di una classificazione multiclasse, nel nostro caso 3, relative alle 3 zone in cui sono stati collocati i microfoni. La cardinalità delle classi vede un 33% di presenza per ognuno.

l'obiettivo è individuare il luogo