

- PR-1.2 G/N. Un problema binario che consiste nel discriminare le due fasi astronomiche principali del giorno, classe G e della notte, classe N. La fascia oraria di riferimento del giorno si presenta tra le ore 6 e le 17, estremi compresi. Le ore restanti sono assegnate alla rispettiva classe che identifica la notte. Le classi presentano un rapporto tra le quantità di circa il 50%. *Considerando G con 17 ore, abbiamo considerato*

- PR-1.3 AT/R. Un tipo di classificazione binaria che considera in un classe i suoni presenti all'alba e al tramonto, classe AT, e nella seconda le ore rimanenti, classe R. La fascia oraria considerata per alba è stata identificata tra le ore 5 e le ore 7, estremi compresi, per il tramonto invece tra le ore 18 e le 20, estremi compresi. *altro problema studio* L'ipotesi alla base di questo problema *perlo studio di* riteneva *è* che nella prima classe *ci possono* potessero esserci dei suoni caratterizzanti e simili rispetto al resto della giornata, riconoscibili nell'alba come il risveglio della natura e nel tramonto come il calare della quiete. Quinn *et alii* evidenziano l'attività della biofonia in tali fasce della giornata [1]. In questa casistica la distribuzione si trova in parte sbilanciata sulla classe R, presente per un 70%.

- PR-1.4 A/T/G/N. Si tratta di una classificazione multiclasse, composta da 4 classi, discriminando alba, classe A, tramonto, classe T, e le ore rimanenti diurne e notturne, classi G e N. Le fasce orarie per l'alba e il tramonto sono le medesime presentate sopra. *Vediamo* le classi A e T in una percentuale di distribuzione minore, con circa un 15% cadauna, rispetto alle classi G e N, che presentano un valore di 40% e 30%. *sono*

- PR-1.5 M. Problema a 3 classi, che *mina e identificare* distingue i relativi mesi in cui sono stati registrati i dati. Per il terzo mese, maggio, è presente un forte squilibrio dovuto alla mancanza di dati, che si identifica con solo un 4% dei dati, rispetto a marzo e aprile con il 49% e 47%.

- PR-1.6 MM. Classificazione binaria che discrimina *Problema* i mesi *mina e* in due parti, la prima *pe* (da inizio mese fino al quindicesimo giorno), la seconda *del mese* (dal sedicesimo giorno fino a fine mese). La distribuzione è ottimamente bilanciata con circa il 50% per entrambe le classi.

In alcune casistiche i dati presentano parti sbilanciate che potrebbero inficiare sulla qualità del risultato: riguarda maggiormente il problema PR1.5 che soffre di una componente poco rappresentata, il mese di maggio. Pur essendo a conoscenza del problema, non si è potuto prendere in considerazione delle tecniche conosciute in letteratura per compensare il divario. Si sottolinea tuttavia, che in tale situazione i risultati ottenuti presentano comunque una valenza considerevole.

Direi: in questo lavoro non è stato considerato, ma sarebbe interessante valutare questo (oggetto di studio futuro)

4.3.2 Rappresentazione con etichette semantiche

La seconda fase utilizza il *dataset* DATA2, sottoinsieme del *dataset* principale. Come precedentemente descritto nel capitolo 3, riducendo il set a un numero di dati censibile si è potuto procedere con una rilevazione dei suoni principali classificando i relativi gruppi. ~~Con una conoscenza più specifica dell'oggetto si è potuto caratterizzare lo studio mediante delle rappresentazioni basate su etichette semantiche, considerando quindi il contenuto, invece che il contesto (come nella prima fase).~~

Sono stati disegnati tre gruppi di problemi di classificazione, suddivisi in base alle classi di cui sono composti:

- Gruppo PR-2.1. Consiste in una classificazione binaria positivo/negativo, ovvero il sistema distingue la presenza o l'assenza di una determinata classe, ~~mediante le etichette con i valori 1 e 0.~~ In questo gruppo sono stati ~~descritti~~ ^{formulati} quattro problemi, con il relativo rapporto percentuale di distribuzione:
 - PR-2.1.1 V. ~~Riguarda~~ ^{identifica} la classe V del veicolo, in rapporto 42/58;
 - PR-2.1.2 G. ~~Identifica~~ ^{new} il suono G dei grilli, con 32/68;
 - PR-2.1.3 P. ~~Identifica~~ ^{trova} la classe P che caratterizza la pioggia, con un bilanciamento di 37/63;
 - PR-2.1.4 T. ~~Riguarda~~ ^{identifica} la classe T che definisce i tuoni, in rapporto 34/66;
- Gruppo PR-2.2. ~~Concerne una classificazione binaria tra due classi distinte, con etichette 1 e 2.~~ Nel seguente gruppo sono stati disegnati i problemi:
 - PR-2.2.1 V/P. Vengono messe in corrispondenza ANT/GEO, con le due classi V e P, in rapporto 55/45;
 - PR-2.2.2 V/G. Si confronta ANT/BIO, con le classi V e G, in rapporto 60/40.
 - PR-2.2.3 G/T. Si relaziona BIO/GEO, con le classi G e T, in rapporto 48/52.
 - PR-2.2.4 G/P. In una seconda contrapposizione BIO/GEO, con le classi G e P, in rapporto 46/54.
- Gruppo PR-2.3. Si tratta di una classificazione ~~ternaria, coinvolge quindi tre classi distinte, identificate da tre etichette, 1, 2 e 3.~~ Verranno descritti i seguenti problemi che mettono in relazione ANT/BIO/GEO:
 - PR-2.3.1 V/G/P. Relazionando le classi V, G e P, in un rapporto di 35/30/35;
 - PR-2.3.2 V/G/T. Relazionando le classi V, G e T, in un rapporto di 32/30/38;

Si sottolinea che la circostanza del ^{il} problema è multi-etichetta, ovvero che ogni audio è caratterizzato da più suoni, quindi più etichette assegnate allo stesso oggetto. Pertanto, per

non
in
2° del
modello

è stato necessario

proseguire con la stessa metodologia applicata nella prima fase e poter disegnare dei problemi di classificazione, per i gruppi PR-2 e PR-3, ~~vi era la necessità di assegnare ad ogni classe solo gli oggetti che la presentassero univocamente all'interno del problema~~: per esempio in una classificazione binaria, in cui vi sono due classi distinte, non è possibile ~~accettare nel dataset un audio con entrambe le etichette, poiché sarebbe assegnabile ad entrambi~~. Dunque sono state ~~definite~~ le combinazioni delle classi dove vi era una distribuzione accettabile ~~dei dati~~ e derivare così una casistica interessante da ~~sottoporre~~. Seguendo questa considerazione, alcune ~~etichette~~ sono state escluse da tutti i problemi a causa di una distribuzione non accettabile, poiché ~~si sarebbe creato uno sbilanciamento in grado di falsare il risultato~~.

includere nel problema le continue entità, i suoi
delle don
per de
con
quel no

- le etichette C e U, relative al suono del fiume/cascata e degli uccelli, si trovano in una percentuale molto alta, di circa il 94% e l'81%;
- le etichette I e S, relative invece alle interferenze e ai suoni non riconosciuti, diversamente, presentano un valore troppo poco rappresentativo, del 3% e 5%.

1 frase introduttiva:

le sequenze sono espresse in un contesto
reale molto complesso, che comporta
la presenza di rumore

abbiamo pensato di provare a vedere
se migliori qualcosa filtrando.

del + alcuni usi
nuovi e poco comuni

4.4 Valutazione del filtraggio di frequenze

In questo paragrafo sarà illustrata la valutazione preliminare effettuata per esplorare l'effetto di un possibile filtraggio delle frequenze dello spettrogramma sul contesto dello studio. Il filtro consiste nell'esclusione delle componenti superiori o inferiori a determinate soglie di frequenza dallo spettrogramma, rendendo più chiare e definite le rimanenti. Nel dettaglio, l'approccio applica filtri passa-basso e passa-alto, che rispettivamente accettano solo determinate frequenze sotto o sopra un limite definito. Applicandoli poi in combinazione si ottiene un filtro passa-banda, che lascia passare delle frequenze intermedie, filtrando in sintonia le superiori e inferiori. Con questi metodi, si dovrebbero eliminare elementi che potrebbero essere causa di rumore, o che, più semplicemente, risultano irrilevanti per lo studio. Tale effetto determina un miglioramento della qualità dei dati, e di conseguenza, perfeziona la capacità discriminativa del classificatore.

Come metodo di valutazione è stato impiegato lo stesso processo utilizzato per la classificazione, quindi valutando la percentuale di errore del modello, alimentato però con i dati filtrati, studiando poi il comportamento di tutte le features nelle due finestre considerate.

Lo studio è stato condotto mediante etichette note del caso PR-1.1 YAT, ma con un dataset più ristretto: si è considerato il mese di marzo, con sei audio al giorno ogni quattro ore, a partire dalle due del mattino, per un totale di 556 file. Si è proceduto in maniera graduale applicando varie combinazioni di filtri, aumentando progressivamente prima dalla parte superiore, viceversa poi dalla parte inferiore, proseguendo in convergenza.

L'esito verrà poi esposto nel capitolo dei risultati, tuttavia si accenna che non è stato osservato nessun miglioramento. Pertanto, la classificazione è stata poi condotta senza l'utilizzo di filtri.

Abbiamo risolto un altro problema, con una
configurazione delle features. In pratica

4.5 Risultati

Prima di esporre e analizzare i risultati ottenuti nei problemi di classificazione, si discuterà l'esito del tentativo di filtraggio delle frequenze ~~condotto~~. Sebbene non abbia portato un effetto ~~applicabile alle casistiche sottoposte~~, è possibile comunque trarre delle conclusioni interessanti.

4.5.1 Filtraggio delle frequenze

Come descritto precedentemente, sono state testate varie combinazioni di filtri utilizzando come metodo di valutazione il risultato della classificazione stessa. Data la quantità di risultati ottenuti si è proceduto sintetizzando con l'esposizione della media matematica su tutte le *features*, come descritto nel paragrafo 4.1 della classificazione.

Nei grafici 4.1 e 4.2 è possibile vedere il risultato del filtraggio. Sull'asse delle ascisse si ~~presenta~~ il valore in frequenza utilizzato per filtrare le due zone dello spettrogramma, nelle varie combinazioni, con un ordinamento diverso tra i due grafici: il primo per filtro inferiore ascendente, il secondo per il superiore. Sull'asse delle ordinate è esposta la percentuale di errore del classificatore. Per potere evidenziare la differenza rispetto alla versione priva di filtraggio, nel grafico si è aggiunta la voce corrisponde 0-0.

Come si evince la differenza tra il caso non filtrato e quelli filtrati è notevole, e tale contrasto denota un peggioramento significativo del risultato all'aumentare dei filtri impiegati.

La frequenza che ottiene i risultati migliori, indifferentemente dal filtro, è la FS1.STD: dal secondo grafico è possibile notare la combinazione di filtri migliore del test derivata dall'utilizzo del solo filtro inferiore nei valori 5, 10 e 15; all'opposto invece la frequenza FS0X.STD ottiene i valori migliori sempre con gli stessi valori, ma attivi sul filtro superiore (visibile nel grafico 1). La versione FS0X.NOR ottiene i risultati peggiori, distaccandosi di più di un centesimo dalle versioni FS0X.STD e la FS1.NOR.

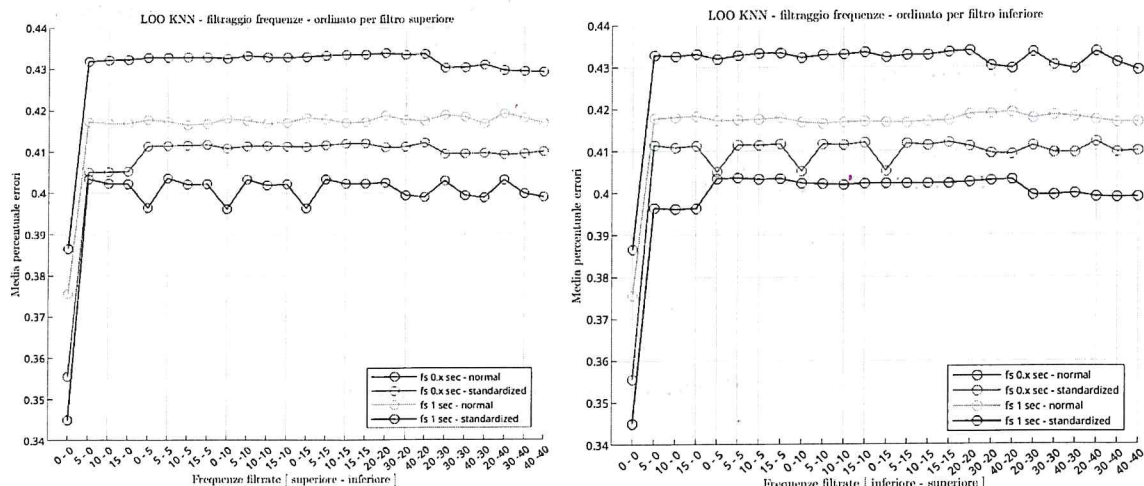


Grafico 4.1 e 4.2 : *metti me descrittore (caption)*

4.5.2 Risultati rappresentazione con etichette note

In questa sezione si riportano i risultati della prima fase di classificazione, basata su rappresentazione tramite categorie note. Per questo studio è giusto tenere conto che il risultato è molto sensibile al fattore considerato nel disegno del problema ovvero il luogo o il tempo.

Nel grafico 4.3 si può visualizzare l'esito dei casi sottoposti: sull'asse delle ordinate troviamo l'errore medio ottenuto dal classificatore, su quello delle ascisse il tipo di problema. Il miglior risultato lo si individua nel caso *PR-1.1 YAT*. Tale caso è l'unico tra quelli con categorie note che tratta della discriminazione del luogo. Sicuramente ha goduto, oltre alle peculiarità della zona in sé, persino delle caratteristiche dello stato dello strumento di misurazione. Si può ipotizzare che elementi come le interferenze o le condizioni meteorologiche, colpendo indifferentemente i vari sensori, abbiano marchiato il prodotto segnando maggiormente uno tra questi rispetto agli altri per un determinato periodo temporale. Quindi hanno caratterizzato una parte del registrato e contribuito ad accrescere le già notabili differenze, definite dall'ambiente circostante, con il resto dei dispositivi. Non di meno, un altro dettaglio considerabile riguarda il suono rappresentante dell'elemento C, che è quasi onnipresente, poiché lo si ritrova nella totalità delle registrazioni, ma in versioni eterogenee tra i dispositivi, dato che sono posizionati a distanze differenti dalla sorgente del rumore. E' possibile affermare che una discriminazione basata sul luogo ha sicuramente beneficiato di queste condizioni rispetto alle altre casistiche.

*cost
voci
due?*

1 frase x ricordare cos'è

*approp
cos
suo
le 4
barre*

*cost
voci
due?*

(riprendi)

D'altra parte, anche le tipologie sviluppate sulla temporalità hanno ottenuto complessivamente dei buoni risultati. L'idea alla base dei casi PR-1.2 G/N e PR-1.3 AT/R di valorizzare i periodi astronomici dell'alba e del tramonto, ha portato delle conclusioni interessanti, in particolare nel secondo, che si basava proprio sul discriminare questi due

momenti principali rispetto al resto della giornata. Pur sapendo che in un contesto naturale

presentano delle caratteristiche sonore uniche, si rende noto, che, diversamente dalle aspettative, non è stato così efficace per il problema PR-1.4 A/T/G/N, risultato con la maggiore percentuale di errori. A sua discolpa si può ipotizzare che il numero delle classi,

quattro in questa situazione, superiore rispetto agli altri, ha sicuramente influito sulla prestazione del modello. Questo può determinare una notevole complessità quindi una

minore capacità di generalizzazione. Inoltre può generare un aumento della variabilità all'interno dei dati con conseguente difficoltà per il modello nel discriminare le classi. Infine,

è possibile sostenere che il risultato sia stato determinato anche dalla riduzione dei dati di addestramento da cui estrarre un modello identificativo della classe stessa: avendo un numero finito di campioni all'aumentare della classi diminuisce il numero di oggetti per descriverle.

Di pari importanza, la capacità delle finestre considerate, si evidenzia dal grafico, mediante i colori delle barre, come la FS1.STD abbia ottenuto le migliori prestazioni in ogni sfida. Si

può ipotizzare che la finestra più ampia riesca a cogliere maggiori dettagli di interesse nel contesto. Una finestra con maggiore risoluzione potrebbe risultare vantaggiosa per

lo studio. A tale risoluzione anche il rumore potrebbe risultare meno influente, non facendo quindi risaltare determinati elementi irrilevanti per la classificazione. Allo stesso modo,

anche l'altra versione standardizzata FS0X.STD ha riportato un esito accettabile e abbastanza simile alla finestra appena descritta. Entrambe le versioni standardizzate hanno conseguito un errore minore rispetto alle controparti regolari. Si considera però che la finestra più corta, ma

con dati standardizzati, FS0X.STD, nella maggioranza dei casi ha ottenuto dei risultati migliori della finestra più ampia ma con dati normali, FS1S.NOR. Quindi si può dedurre che

per il contesto la scelta corretta sia una combinazione tra la finestra più ampia e dati standardizzati.

La medesima situazione è osservabile sotto un'altra prospettiva offerta dal grafico 4.4, dal punto di vista delle features, ottenuta con le medie dei valori trasversalmente alle casistiche.

Si osserva che nei gruppi con le features singole ORIG le capacità per tipologia di finestra si discostano di poco, diversamente che nelle forme combinate dove le versioni standardizzate sono maggiormente incisive. Lo si può notare nel gruppo CONC.MED.ORIG, che ottiene i

migliori risultati. Riassume le caratteristiche essenziali dei tre gruppi SPE/TON/TEM,

metti
numeri
e esempi
(come
prima)

non
sono
d'accordo
le quantità

probabile di
confusione
2000

si può
osservare
come

idea
spiega
bene

ma
può
numero x
capire meglio

servendosi di un numero ristretto di *features*, undici in questo caso. Un insieme di componenti ridotto, riduce la complessità del modello, rendendolo più robusto, aumentando la sua interpretabilità e ottenendo una migliore generalizzazione.

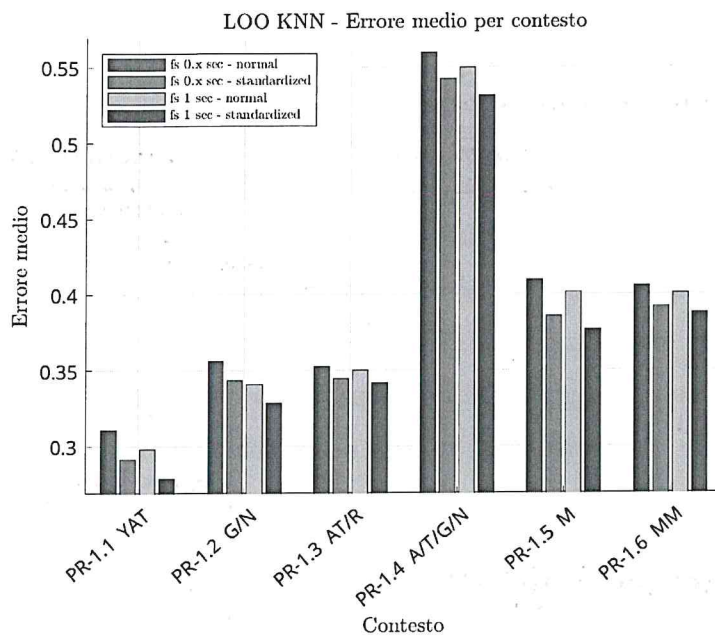


Grafico 4.3 : mostra la decisione

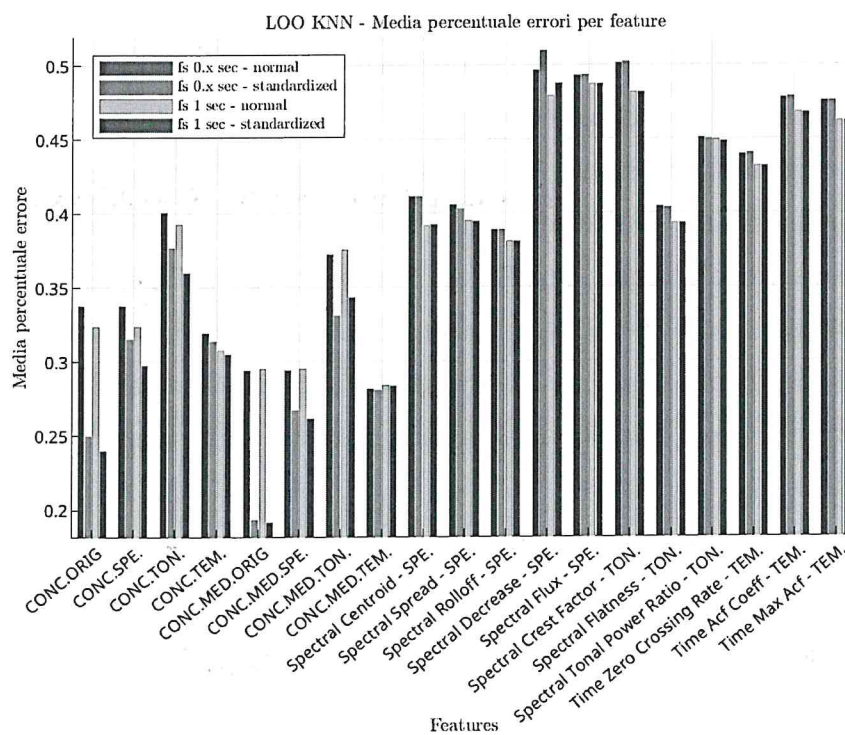


Grafico 4.4 anche per