24/03/2021

**Considerazioni sui risultati di Giulia e Francesco sugli eventi dal piano Galattico**

***Nota:*** *i numeri entro parentesi ( ) si riferiscono a figure/tabelle/eq dell’elenco puntato.*

1. Assumiamo la stessa regione angolare che abbiamo assunto nell’articolo [1],

|  |  |
| --- | --- |
| La regione ON corrisponde alla banda |b|<3o e |l|<40o, per un angolo solido =0.145 sr.  Figura: proiezione Aitoff della regione ON e di quelle OFF per il calcolo del fondo |  |

1. Immaginiamo ora di voler usare gli eventi selezionati da Federico nell’articolo sui neutrini atmosferici [2]. Quanti eventi di segnale/fondo possiamo aspettarci? Ovviamente, dipende dal taglio sulla variabile BDT Federico (BDT).

|  |  |
| --- | --- |
|  |  |
| Distribuzione BDT per muoni atmosferici e neutrini atmosferici/cosmici. La tabella a sinistra (Tab. 1 di [2]) mostra gli eventi dopo la preselezione, ossia tutti quelli presenti in Figura | |

1. Con il taglio BDT>0.33, usato in [2] tutti i CRmu sono eliminati. In particolare, rimangono

|  |  |  |
| --- | --- | --- |
| * **numu\_CC= 620** nu\_mu “tracce” * **Not\_numu\_CC= 285** (96+9+180) eventi “cascata” * circa 9 eventi “cosmici” con il flusso dato dalla formula: |  |  |

1. Giulia, usando la figura (2), stima che un taglio BDT>0.12 porterebbe a una contaminazione del campione da parte di CRmu pari a circa il 14%. Questo valore è stato scelto in maniera che sia analogo a quello presente nel campione dei mu passanti. La tabella riporta il numero di eventi per i due tagli, assumendo il flusso di neutrini atmosferici (i numeri per BDT>0.12 provengono da Francesco). Notare che diminuendo ulteriormente il valore di BDT rispetto al valore 0.12, il numero di Not\_numu\_CC (distribuzione “rossa”) non aumenterebbe di molto.

|  |  |  |  |
| --- | --- | --- | --- |
| **Tipo eventi** | **No BDT cut** | **BDT>0.12** | **BDT>0.33** |
| Not\_Numu\_CC | **664** | 557 | 287 |
| Numu\_CC | **3780** | 2758 | 626 |
| CRmu | **136700** | 530 | 0 |
| TOTALE |  | 3315+530 | 913 |

1. Il punto fondamentale trovato da Giulia riguarda la risoluzione angolare degli eventi selezionati. La figura sotto riporta la “risoluzione angolare” TANTRA vs AAFIT per gli eventi con BDT>0.12. Circa il 50% degli eventi Not\_numu\_CC sono ricostruiti entro 3o dalla direzione vera con TANTRA. Per i numu\_CC invece tale metodo fornisce precisioni inaccettabili. Anche AAFIT non fornisce valori adeguati (circa il 50% ricostruito entro 5o). Il problema è quindi: **solo gli eventi Not\_numu\_CC sono ricostruite con precisione sufficiente da rientrare (con probabilità sufficiente) nella regione ON definita nel punto (1). Occorre trovare un metodo per discriminare (ed eliminare) gli eventi numu\_cc** .

|  |
| --- |
|  |
| Eventi Distribuzione cumulativa di eventi (BDT>0.12 e pesati con spettro E^-2.3) con differenza angolare “|direzione vera-direzione ricostruita|” inferiore al valore in asse x. La linea tratteggiata verde corrisponde al 50% degli eventi. Le linee blu rappresentano il risultato di TANTRA, quelle rosse quello di AAFIT. A sinistra, “numu\_CC”, a destra gli altri “Not\_numu\_CC” |

1. **Stima (rozza!) degli eventi attesi nella regione di segnale ON**. Per questo possiamo usare la stima dall’ICRC [3]



I 9.2 eventi della Tab. del punto (3) corrispondono a tale flusso, integrato sui 2pi sr degli eventi upgoing. Possiamo far finta che tale flusso corrisponda al segnale ON dalla nostra regione, ossia concentrato nella regione =0.145 sr di (1). Ciò corrisponde a un flusso (1flavor) 1.5 10-18 (E/100 TeV)-2.3 [GeV cm2 s]. Proviamo a disegnare tale flusso nella figura col limite ottenuto in [1]. Si ottiene la linea marrone sotto riportata

|  |  |
| --- | --- |
| Limite ottenuto nella regione ON in [1]. Per stimare il numero di eventi attesi, usando la simulazione di Federico per il flusso cosmici diffusi, facciamo finta che tutto il segnale sia concentrato nella regione ON.  La formula 3.3 di [3] sopra riportata darebbe  F1f(100 TeV)=1.5 10-8 (E/100 TeV)-0.3 GeV cm-2 s-1,  Ossia:  F1f(1 GeV)= 4.7 10-7 GeV cm-2 s-1,  F1f(10 GeV)= 2.4 10-7 GeV cm-2 s-1.  La figura marrone è per 3 flavors, quindi I numeri riportati sopra sono x3.  Si noti che la curva è sotto il limite di ANTARES e vicino al modello KRA. Può essere considerato quindi un buon stimatore del numero di eventi attesi nella regione ON |  |

1. Stima di eventi “segnale” con i tre valori di taglio di BDT (valori per canale da Francesco). Col taglio >0.12 il numero di eventi cosmici raddoppia, mentre i neutrini atmosferici triplicano (4). Senza il cut di BDT, ho stimato io il numero con (?). La somma invece è data nella tabella (2).

|  |  |  |  |
| --- | --- | --- | --- |
| **Tipo eventi** | **No BDT cut** | **BDT>0.12** | **BDT>0.33** |
| Not\_Numu\_CC | 17 (? stima) | **11.8** | **6.8** |
| Numu\_CC | 13.4 (? stima) | **9.8** | **2.7** |
| TOTALE | **30.4** | **21.6** | **9.5** |

1. Francesco ha sviluppato un algoritmo basato su neural network (chiamiamolo NNFra) con lo scopo di separare i due campioni di Numu\_CC da quelli di Not\_Numu\_CC. In base alle considerazioni sulla risoluzione angolare in (5), gli eventi Numu\_CC sono “cattivi”, in quanto non ricostruiti da TANTRA (e anche da AAFIT!) con sufficiente precisione angolare.

NNFra ha diverse variabili, ma quelle che sembrano più caratterizzanti i due campioni sono in numero di hit on time, trigger counter e tantra hits.

1. Nella riunione “KM3-Bo” del 19/3 si è discusso un po’ di nomenclatura. Francesco ha chiamato “0” gli eventi Not\_Numu\_CC (P, positivi, in letteratura) “1” gli eventi Numu\_CC (N, negativi). La NNFra può riconoscerli veri (T) o falsi (F), potendosi avere la seguente tabella di verità:

|  |  |  |
| --- | --- | --- |
| NN “Positivi” | **TP** | FP |
| NN “Negativi” | FN | **TN** |
|  | Not\_Numu\_CC | Numu\_CC |

Lo scopo dell’algoritmo sarebbe quello di avere gli elementi diagonali =1 e gli elementi non diagonali =0. In letteratura (medicina, ingegneria), vengono definite le seguenti variabili. Francesco, come parametro di controllo, utilizza l’accuracy.

|  |  |
| --- | --- |
| **Sensitivity = TP/(TP+FN)** | Ossia, la percentuale degli eventi ricostruiti come “buoni” rispetto al totale di quelli che avremmo voluto accettare |
| **Specificity = TN/(TN+FP)** | Analogo alla “sensitivity”, per gli eventi che avremmo voluto buttare |
| **Accuracy = (TP+TN)/(TN+TP+FN+FP)** | Rapporto tra eventi ricostruiti correttamente, rispetto al totale |
| **Purity = TP/(TP+FP)** | Percentuale degli eventi ricostruiti come “buoni” rispetto al totale di quelli che accettiamo |

1. Francesco ha addestrato una NN particolarmente efficiente usando tutti gli eventi con BDT>0.33 e poi con BDT>0.12 ma non pensandoli per alcuno spettro energetico. A sinistra, “permutation importance” dell’algoritmo. A destra, la classificazione: estremamente incoraggiante.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Numeri in migliaia di eventi (percentuale)  **BDT>0.33**   |  |  | | --- | --- | | 544 (90%) | 56 | | 58 | 741 (92%) | |

**Problema: può questo risultato incoraggiante dipendere da un bias dovuto al fatto che gli eventi non sono stati pesati per un particolare spettro? Fare le tabelle di verità in diversi range di energia.**

1. Analisi della NNFra in diversi intervalli di energia. Lo spettro di energia per i due campioni appare come in figura (BDT>0.12). Ho suggerito 5 intervalli che contenessero un numero equivalente di eventi non pesati. Il risultato è mostrato a destra: la curva rossa rappresenta gli eventi TN e quella blu i TP (ossia, gli eventi lungo la diagonale). Il risultato è estremamente incoraggiante: **la NN riconosce con la stessa efficienza eventi che differiscono di energia vera per diversi ordini di grandezza. Non occorre addestrare la NN per uno specifico indice spettrale/modello. Risultato model-independemt!**

|  |  |
| --- | --- |
|  |  |

1. Come viene definito il taglio per classificare gli eventi? Il risultato della NNFra è un numero compreso tra 0 e 1, come si vede nella figura sotto riportata.

|  |  |
| --- | --- |
| Considerazioni: il taglio viene posto a 0.5. Gli eventi che vogliamo accettare sono gli 0. Potrebbero essere aumentati prendendo NNout più grande, ma la scala logaritmica mostra che questi sarebbero poco di più. Tutto sommato, sembra una scelta adeguata, forse non è necessario un lavoro di ottimizzazione. |  |

1. Consideriamo ora i risultati di Francesco PESATI per lo spettro atmosferico e per cosmico (per BDT>0.33, dobbiamo riottenere i risultati dell’articolo [2] e della tabella (3)

|  |  |  |
| --- | --- | --- |
| BDT | ATMOSFERICI | COSMICI |
| >0.33 | |  |  |  | | --- | --- | --- | | NN”0” | NN”1” |  | | 270 | 17 | (287) | | 108 | 518 | (626) | | **378** | **535** | (913) | | |  |  |  | | --- | --- | --- | | NN”0” | NN”1” |  | | 6.6 | 0.2 | (6.8) | | 0.9 | 1.6 | (2.7) | | **7.5** | **1.8** | **(9.5)** | |
| >0.12 | |  |  |  | | --- | --- | --- | | NN”0” | NN”1” |  | | 499 | 58 | (557) | | 241 | 2517 | (2758) | | **740** | **2575** | (3315) | | |  |  |  | | --- | --- | --- | | NN”0” | NN”1” |  | | 11.2 | 0.6 | (11.8) | | 2.0 | 7.9 | (9.9) | | **13.2** | **8.5** | **(21.7)** | |
| CRmu  >0.12 | Per BDT>0.33 non ci sono CRmu   |  |  |  | | --- | --- | --- | | NN”0” | NN”1” |  | | 0 | 0 | (0) | | 31.2 | 500 | (531) | | **31** | **500** | (531) | |  |

1. Cerchiamo di stimare l’effetto di segnale/fondo nella regione ON, tenendo conto delle risoluzioni angolari degli eventi ricostruiti con TANTRA o AAFIT (5) e classificati dalla NN (13). La regione di angolo solido =0.145 sr in (1) corrisponde a **0.145/2pi=0.023** del totale degli eventi di fondo (assumendo una efficienza di ricostruzione identica nell’angolo solido. Quindi:

* Gli eventi di fondo (ATMO e CRmu) vanno quindi moltiplicati per 0.023.

Invece per gli eventi di segnale, assumiamo usando le risoluzioni angolari di (5) che con il taglio |b|<3 gradi vengano accettati

* il 50% degli eventi TP (ossia, eventi Not\_Numu\_CC ricostruiti positivi) usando TANTRA,
* il 40% degli eventi TN Numu\_CC ricostruiti negativi, usando però AAFIT
* 0% degli eventi FN e FP (la precisione angolare di TANTRA per loro è pessima);

|  |  |
| --- | --- |
| Segnale BDT>0.33 | Fondo BDT>0.33 |
| |  |  | | --- | --- | | NN”0” | NN”1” | | 6.6x0.5 | 0.2x0 | | 0.9x0 | 1.6x0.4 | | **3.3** | **0.6** | | |  |  | | --- | --- | | Fondo Atmo | | | NN”0” | NN”1” | | **378** | **535** | | X0.023 | X0.023 | | **8.7** | **12.3** | |

|  |  |
| --- | --- |
| Segnale BDT>0.12 | Fondo BDT>0.12 |
| |  |  | | --- | --- | | NN”0” | NN”1” | | 11.2x0.5 | 0.6x0 | | 2.0x0 | 7.9x0.4 | | **5.6** | **3.1** | | |  |  | | --- | --- | | Fondo Atmo | | | NN”0” | NN”1” | | **740+30** | **2575** | | X0.023 | X0.023 | | **17.7** | **59** | |

Mi sembra che l’opzione di considerare anche gli NN”1”(ossia la colonna rossa), sia da scartare (il segnale aumenta di poco, il fondo cresce moltissimo). Nella stima grossolana abbiamo

* **BDT>0.33 , nella regione ON sono attesi 8.7 eventi di fondo, e 3.3 eventi di segnale;**
* **BDT>0.12 , nella regione ON sono attesi 17.7 eventi di fondo, e 5.6 eventi di segnale;**

**Referenze**

[1] Galactic ridge. Physics Letters B760 (2016) 143.

[2] Atmospheric neutrinos. Physics Letters B816 (2021) 136228

[3] Fusco PoS(ICRC2019)891