

Katalóg požiadaviek

Rekonštrukcia top kvarkov

Richard Mészáros

Martin Slimák

Magdaléna Kozubaľová

Veronika Benková

Obsah

1. Úvod.....	3
1.1 Účel katalógu požiadaviek.....	3
1.2 Rozsah využitia systému.....	3
1.3 Slovník pojmov	3
1.4 Odkazy a referencie	4
1.5 Prehľad nasledujúcich kapitol	4
2. Všeobecný popis	5
2.1 Perspektíva systému	5
2.2 Funkcie systému	5
2.3 Charakteristika používateľa	5
2.4 Všeobecné obmedzenia	6
2.5 Predpoklady a závislosti	6
3. Špecifické požiadavky	7
3.1 Funkčné požiadavky	7
3.2 Kvalitatívne požiadavky	8
3.3 Požiadavky rozhrania	8

1. Úvod

1.1 Účel katalógu požiadaviek

Tento dokument slúži na špecifikovanie a súhrn požiadaviek od zadávateľa pre aplikáciu rekonštrukcia Top Kvarkov. Bol vytvorený na základe komunikácie so zadávateľom projektu. Je určený pre zadávateľa, užívateľa a kohokoľvek, kto sa chce dozvedieť na čo slúži a kto so systémom bude pracovať. Účel tohto dokumentu je zabezpečenie správnosti a úplnosti dohodnutých požiadaviek od zadávateľa. Taktiež je dokument záväzný pre obe strany.

1.2 Rozsah využitia systému

Cieľom projektu je vytvorenie systému, ktorý bude pomocou neurónovej siete priradovať jety ku kvarku z ktorého vznikol. Ďalším cieľom je porovnanie výstupu z KLfitter-u a neurónových sietí vo webovom rozhraní prostredníctvom dát z databázy. Webové rozhranie bude slúžiť na komunikáciu s databázou, serverom atlas23 a interakciou s používateľom.

1.3 Slovník pojmov

Kvark - podľa štandardného modelu časticovej fyziky elementárne častice, z ktorých sa skladajú hadróny (teda napríklad protóny a neutróny).

Top kvark – je kvark tretej generácie s elektrickým nábojom $+(2/3)e$. Je to najťažší zo všetkých známych elementárnych častíc.

Jet - úzky kužeľ hadrónov a iných častíc produkovaných hadronizáciou kvarku alebo gluónu v experimente s fyzikou častíc alebo experimentom s ťažkými iónmi.

KLfitter – (Kinematic Likelihood Fitter) knižnica na kinematickú montáž pomocou pravdepodobnosti. Je primárne vyvinutý pre prípad rekonštrukcie top kvarku, ale dá sa ľahko upraviť tak, aby vyhovoval aj iným procesom.

Neurónová sieť – výpočtový model, zostavený na základe abstrakcie vlastností biologických nervových systémov.

Hlboká neurónová sieť – výpočtový model, zostavený na základe abstrakcie vlastností biologických nervových systémov s viacerými vrstvami.

Tensorflow – bezplatná a otvorená softvérová knižnica pre dataflow a diferencovateľné programovanie v rôznych úlohách. Je to matematická knižnica a používa sa tiež na aplikácie strojového učenia, ako sú neurónové siete.

KERAS – vysokoúrovňové API pre neurónové siete, schopné bežať na vrchole Tensorflow. Umožňuje rýchle experimentovanie prostredníctvom vysokoúrovňového, ľahko použiteľného, modulárneho a rozšíriteľného API.

ATLAS23 – Server

TTbar event - zrážka top kvakových párov

konfiguračný súbor – súbor obsahujúci všetky potrebné parametre na tréning a následne používanie neurónovej siete.

Procesorový čas - čas potrebný na vykonanie výpočtov na procesore.

1.4 Odkazy a referencie

1.4.1. ROOT Framework na analýzu dát: <https://root.cern.ch/>

1.4.2. Ako čítať strom z ROOT: <https://root.cern.ch/how/how-read-tree>

1.4.3. KL Fitter: <https://github.com/KLFitter/KLFitter>

1.4.4. LWTNN knižnica na prácu s neurónovými sieťami: <https://github.com/lwttn/lwttn>

1.5 Prehľad nasledujúcich kapitol

V nasledujúcich kapitolách sa čitateľ oboznámi s funkciami, perspektívami a obmedzeniami systému, taktiež so špecifickými požiadavkami. Tretia kapitola hovorí o funkčných a kvalitatívnych požiadavkách systému.

2. Všeobecný popis

Implementácia neurónovej siete a optimalizácia metódy KLFitter slúžiacej na rekonštrukciu top kvarkových párov.

2.1 Perspektíva systému

Systém bude ponúkať 2 možnosti výberu rekonštrukcie tt-bar eventov :

1. hlboké neurónové siete
2. KLFitter

Tieto možnosti si bude užívateľ môcť vybrať vo webovom rozhraní. Vo webovom rozhraní bude možné spustiť tréning neurónovej siete, pre ktorú sa budú môcť vo webovom rozhraní nastaviť parametre. Systém bude využívať aj databázu, ktorá bude uložená na serveri ATLAS23. Po spracovaní dát metódami na rekonštrukciu top kvarkov sa výsledky z týchto metód uložia do databázy, aby sme ich mohli porovnávať s nastaveniami, ktoré sme na začiatku zvolili. Výstupom budú dáta v grafickej a textovej podobe, ktoré budú reprezentovať účinnosti použitých metód.

2.2 Funkcie systému

Systém bude ponúkať možnosť vybrať si KLFitter alebo neurónovú sieť na rekonštrukciu top kvarkov. Vstupom bude súbor typu ROOT , ktorý sa prekonvertuje pomocou nástroja na prácu (so súbormi typu ROOT) na vhodný dátový typ, s ktorým vie systém pracovať. V prípade, že si používateľ vyberie neurónovú sieť, systém vytvorí skonvertovaný konfiguračný súbor , ktorý je potrebný pre tréning neurónovej siete. Poslednú natrénovanú neurónovú sieť bude systém vyhodnocovať a následne výsledky uloží do databázy. Výsledky budú v databáze uložené v tabuľke, z ktorej si používateľ môže pomocou webového rozhrania alebo príkazového riadku vybrať všetky požadované údaje z jej histórie napr. dátum, konfiguračný súbor, výsledky, staršie natrénované neurónové siete. Webové rozhranie umožní porovnanie výstupov KLFittera a neurónovej siete alebo porovnanie dvoch neurónových sietí.

2.3 Charakteristika používateľa

Tento systém je určený pre fyzikálnych výskumníkov, ktorí pracujú s top kvarkami a jetmi. Systém bude použiteľný jedine po prihlásení.

2.4 Všeobecné obmedzenia

System bude využívať server, databázu a pripojenie na internet.

2.5 Predpoklady a závislosti

- úložný priestor na serveri
- internet
- server na ktorom bude bežať program

3. Špecifické požiadavky

3.1 Funkčné požiadavky

3.1.1. KIFitter - Systém ponúkne využitie metódy KLFitter na rekonštrukciu top kvarkov. Algoritmus pracuje s permutáciami a ráta pravdepodobnosti zoradenia jetov. Na výpočet pravdepodobnosti táto metóda využíva fyzikálne informácie. Výsledok uloží do databázy.

3.1.2. Hlboká neurónová sieť – Systém ponúkne ďalšiu metódu na rekonštrukciu top kvarkov. Podľa vstupných súborov pre natréňovanie neurónovej siete, systém natrénuje neurónovú sieť podľa konfiguračných súborov obsahujúcich všetky potrebné parametre a následne použije túto metódu.

3.1.4. Databáza - Výsledky z metód KLFitter a neurónovej siete bude systém ukladať do databázy. Okrem uložených výsledkov z metód KLFittera a neurónovej siete, bude obsahovať aj informáciu o dĺžke procesorového času, ktorý bol potrebný na vykonanie metód.

3.1.5. Porovnávanie – Systém dokáže aplikovať metódy KLFitter a neurónovej siete nezávisle od seba. Systém môže porovnať KLFitter s neurónovou sieťou alebo môže porovnať dve neurónové siete, každá s inými parametrami. Výsledky metód sa uložia do databázy. Používateľ si bude môcť vo webovom rozhraní výsledky pozrieť v grafe, porovnať ich.

3.1.6. Spúšťanie - Systém sa bude dať spustiť dvoma spôsobmi. Prvou možnosťou bude spustenie cez príkazový riadok. Druhá možnosť je spustenie cez webové rozhranie, čo znamená, vygenerovanie príkazu pre príkazový riadok.

3.1.7. Premenné - Vo webovom rozhraní si používateľ zvolí premenné ručne. Bude si môcť vybrať z 20 premenných. Ide o premenné, ktoré by štandardne boli vo vstupnom súbore.

3.1.8. Human readable - Výstup systému bude human readable – čitateľný formát napr.: yaml, json, plain text a pod., aby sa výsledky dali zhodnotiť aj ručne.

3.1.9. Vstupné súbory – Systém bude vedieť pracovať so súbormi typu ROOT. Vstupné súbory sa nachádzajú na serveri (atlas23). Obsahujú údaje o permutáciách. Obsahom sú tabuľky (keys). Súbor je rozdelený na 2 tabuľky: signal_permutation a background_permutation. Tieto tabuľky sú typu TTree. Každá TTree je rozdelená na niekoľko stĺpcov typu TBranch. V týchto tabuľkách sa nachádzajú údaje o eventoch, premenných, permutáciách a pod.

3.1.10. Trénovanie - Výstup z tréovania DNN bude použiteľný v C++ (pozrieť bod 1.4.4 - LWTNN [4])

3.2 Kvalitatívne požiadavky

- **Procesy** - Systém by mal vedieť spustiť viac procesov (vstupov) naraz.
- **Stabilita** - Systém by nemal padať.
- **Rýchlosť** - Systém by mal pracovať čo najefektívnejšie, avšak pre veľké vstupy sa doba bežania odhadnúť nedá. (tzn. systém môže bežať niekoľko hodín, dní... ale nevieme to vopred odhadnúť). Aj napriek tomu bude potrebné optimalizovať algoritmy KLFitteera, aby bežal rýchlejšie.

3.3 Požiadavky rozhrania

- **Výber metódy** - V rozhraní sa bude nachádzať možnosť výberu, či chceme robiť rekonštrukciu top kvarkov pomocou neurónovej siete alebo cez KLFlitter
- **Heslo** - Rozhranie bude zaheslované
- **Stav** - Rozhranie bude zobrazovať, na koľko percent je proces dokončený