

Uporaba strojnega učenja pri odkrivanju nove fizike na LHC

Seminar

Simon Bukovšek

Oddelek za fiziko
Fakulteta za matematiko in fiziko
Univerza v Ljubljani

10. april 2024

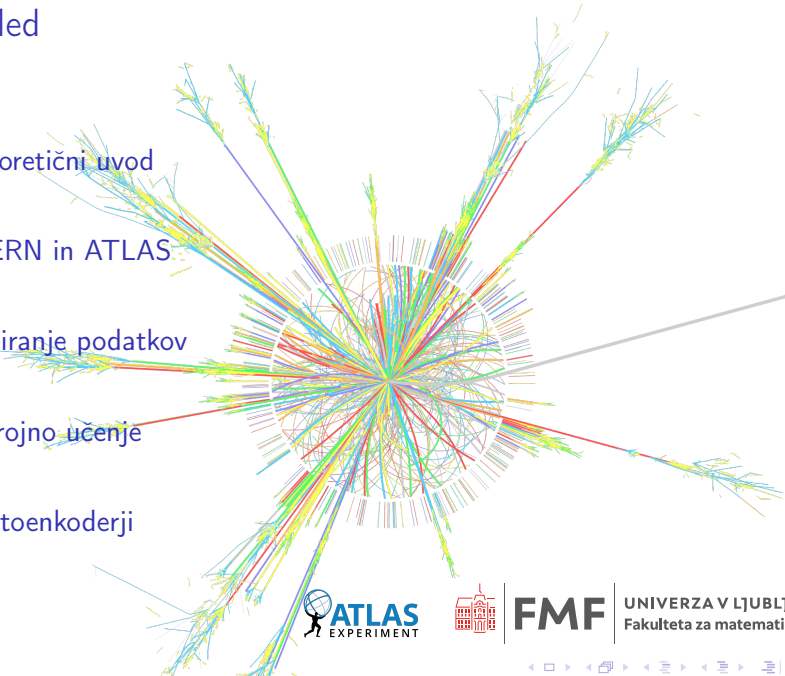


FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Pregled

- 1 Teoretični uvod
- 2 CERN in ATLAS
- 3 Zbiranje podatkov
- 4 Strojno učenje
- 5 Avtoenkoderji



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Standardni model (SM)

- Leptoni, kvarki in umeritveni bozoni
- Polja in simetrije
- Tri vrste interakcij: šibka, močna in elektromagnetna
- Najnatančnejši fizikalni model, kar jih poznamo

		tri generacije fermionov			nosilci interakcij	
		I	II	III		
kvarki	masa	$\approx 2.2 \text{ MeV}$	$\approx 1.28 \text{ GeV}$	$\approx 173,1 \text{ GeV}$	0	$\approx 125 \text{ GeV}$
	naboj	$2/3$	$2/3$	$2/3$	0	0
	spin	$1/2$	$1/2$	$1/2$	1	0
		u gor	c čar	t vrh	g gluon	H Higgsov bozon
		d dol	s čudo	b dno	γ foton	
		e elektron	μ mion	τ taunon	Z bozon Z	
leptoni		ν_e elektronski neutrino	ν_μ mionski neutrino	ν_τ tauonski neutrino	W^\pm bozona W^+ in W^-	umeritveni bozoni

Slika: Standardni model

Standardni model (SM)

		tri generacije fermionov			nosilci interakcij	
		I	II	III		
kvarki	masa naboj spin	$\approx 2,2 \text{ MeV}$ $2/3$ $1/2$	$\approx 1,28 \text{ GeV}$ $2/3$ $1/2$	$\approx 173,1 \text{ GeV}$ $2/3$ $1/2$	0 0 1	$\approx 125 \text{ GeV}$ 0 0
		u gor	c čar	t vrh	g gluon	H Higgsov bozon
		d dol	s čudo	b dno	γ foton	
leptoni		$\approx 0,511 \text{ MeV}$ -1 $1/2$	$\approx 106 \text{ MeV}$ -1 $1/2$	$\approx 1,78 \text{ GeV}$ -1 $1/2$	$\approx 91,2 \text{ GeV}$ 0 1	
		e elektron	μ mion	τ tauon	Z bozon Z	
daljni delci		$\approx 0 \text{ eV}$ 0 $1/2$	$\approx 0 \text{ eV}$ 0 $1/2$	$\approx 0 \text{ eV}$ 0 $1/2$	$\approx 80,4 \text{ GeV}$ ± 1 1	
		ν_e elektronski nevtrino	ν_μ mionski nevtrino	ν_τ tauonski nevtrino	W^\pm bozona W^+ in W^-	umeritveni bozoni

Slika: Standardni model

Pomanjkljivosti Standardnega modela

- Gravitacija
- Materija – antimaterija
- Mase osnovnih delcev
- Temna snov in energija

tri generacije fermionov			nosilci interakcij		
	I	II	III		
masa	$\approx 2.2 \text{ MeV}$	$\approx 1.28 \text{ GeV}$	$\approx 173.1 \text{ GeV}$	0	$\approx 125 \text{ GeV}$
naboj	$2/3$	$2/3$	$2/3$	0	0
spin	$1/2$	$1/2$	$1/2$	1	0
	u gor	c čar	t vrh	g gluon	H Higgsov bozon
kvarki	d dol	s čudo	b dno	γ foton	umeritveni bozoni
	e elektron	μ mion	τ taunon	Z bozon Z	
leptoni	ν_e elektronski neutrino	ν_μ mionski neutrino	ν_τ taunski neutrino	W^\pm bozoni W^+ in W^-	

Slika: Standardni model



CMS Experiment

Evropski center za jedrske raziskave

ATLAS Experiment

ALICE Experiment

LHCb Experiment



FMF

UNIVERSITY OF LJUBLJANA
Fakulteta za matematiko, fiziko
in informatiko



CMS Experiment

Evropski center za jedrske raziskave

Conseil européen pour la Recherche nucléaire

ATLAS Experiment

ALICE Experiment

LHCb Experiment



FMF

UNIVERSITY OF LJUBLJANA
Fakulteta za matematiko, fiziko
in informatiko



CMS Experiment

Evropski center za jedrske raziskave CERN

ATLAS Experiment

ALICE Experiment

LHCb Experiment

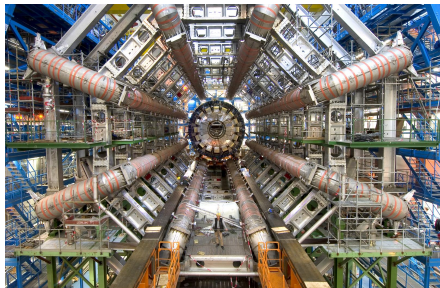


FMF

UNIVERSITY OF LJUBLJANA
Fakulteta za matematiko in fiziko

Veliki hadronski trkalnik (LHC)

- Obseg 27 km
- Dve cevi z gručami protonov
- Razpoložljiva težiščna energija $\sqrt{s} = 13 \text{ TeV}$
- Več eksperimentov, med njimi ATLAS (A Toroidal LHC ApparatuS)



Slika: Ogrodje eksperimenta ATLAS

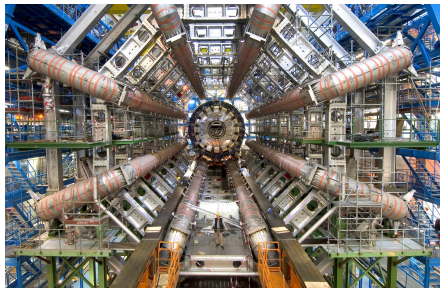


FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Veliki hadronski trkalnik (LHC)

- Obseg 27 km
- Dve cevi z gručami protonov
- Razpoložljiva težiščna energija $\sqrt{s} = 13 \text{ TeV}$
- Več eksperimentov, med njimi ATLAS
(A Toroidal LHC ApparatuS)



Slika: Ogrodje eksperimenta ATLAS

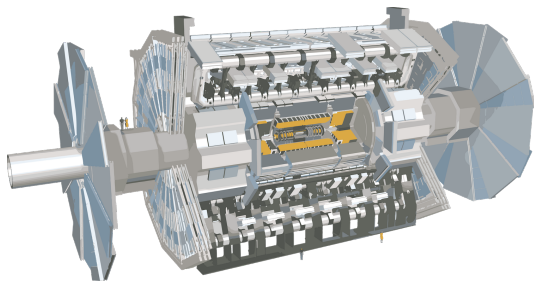


FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

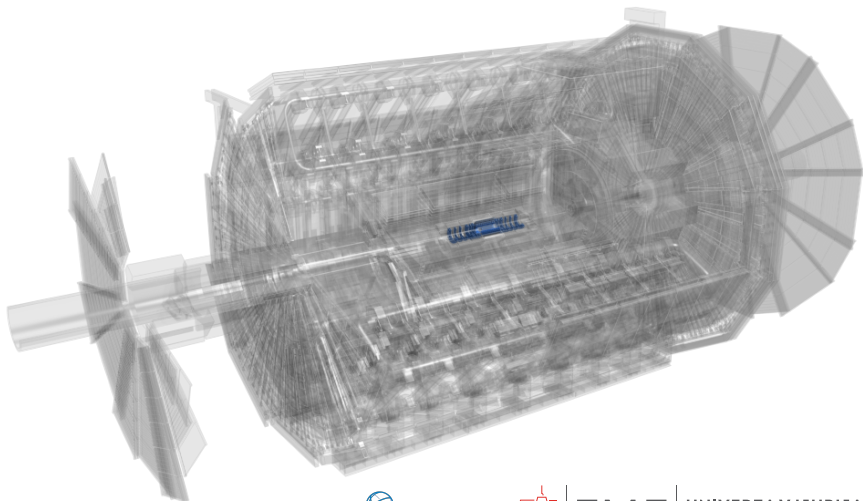
Sestava detektorja ATLAS

- Čebulna struktura
- Detekcija različnih delcev
- Zajem podatkov



Slika: Sestava detektorja ATLAS

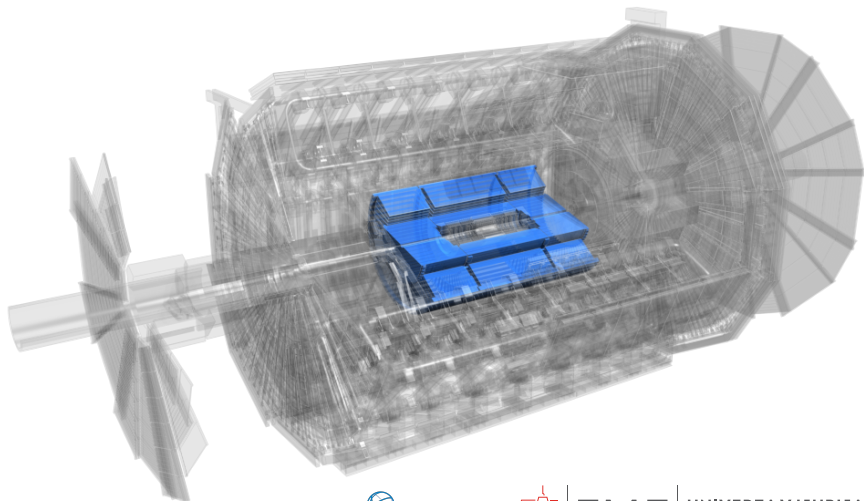
Sledilni detektor



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

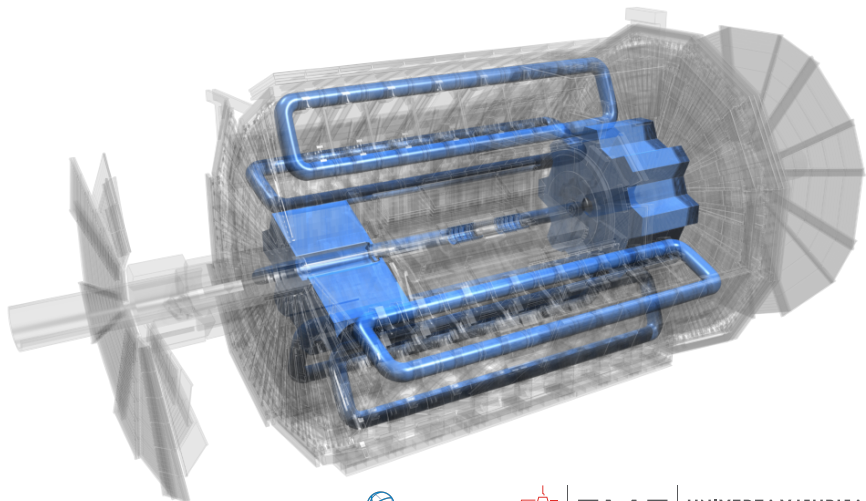
Hadronski in EM kalorimeter



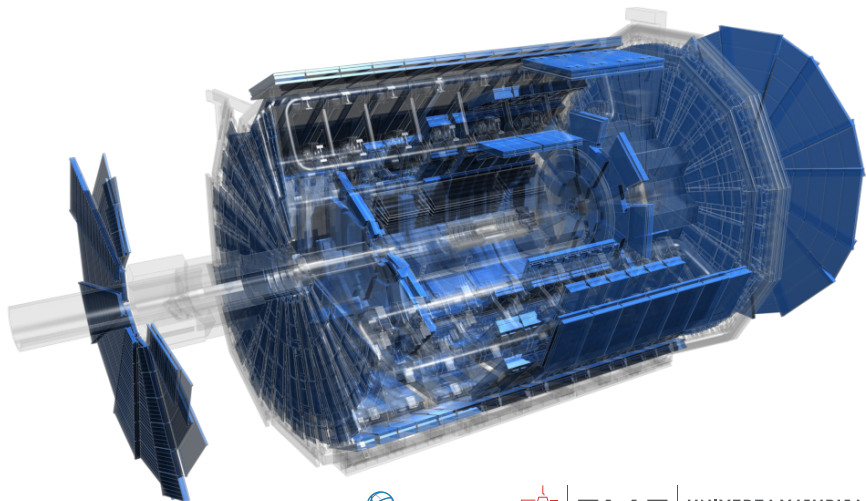
FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Magnetni sistem

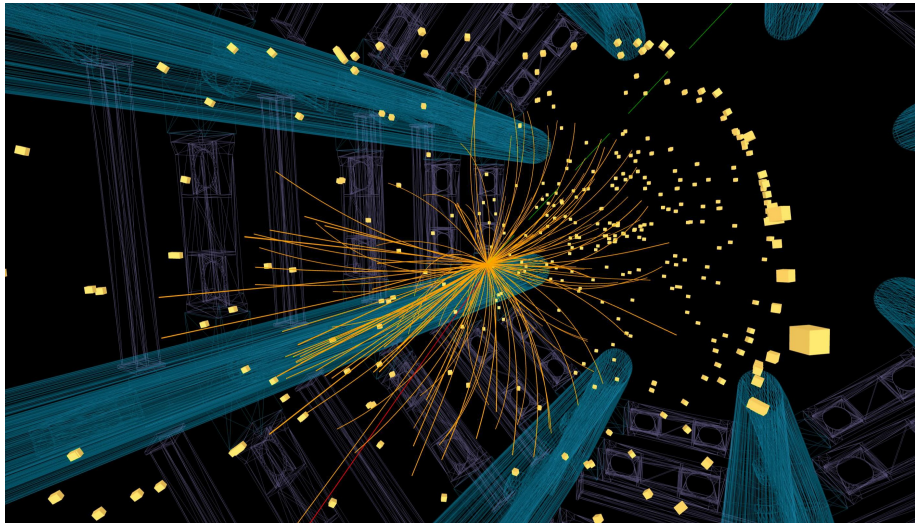


Mionski spektrometer



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

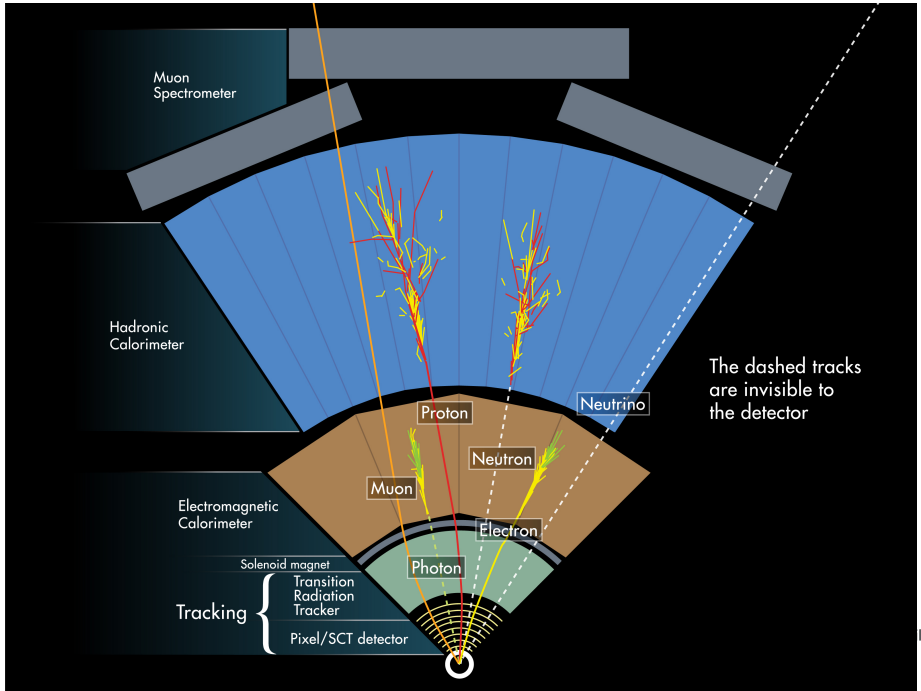


Slika: Primer dogodka v detektorju ATLAS



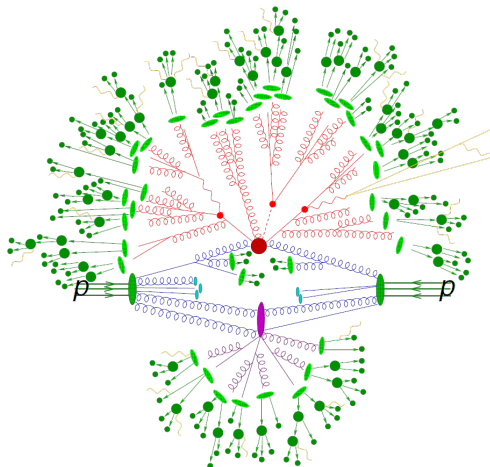
FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko



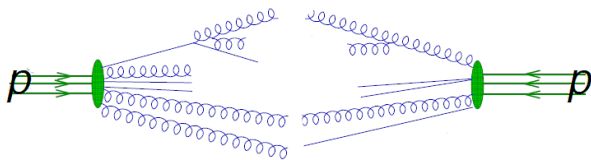
Simulacije dogodkov

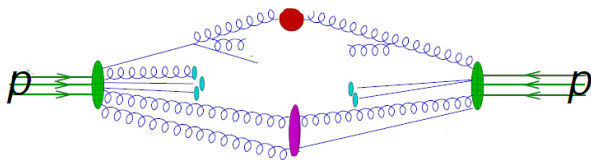
- Monte Carlo simulacije
- Generiranje dogodkov v več korakih
- Simuliranje interakcije med delci in detektorjem
- Lahko simulirajo tudi hipotetične procese

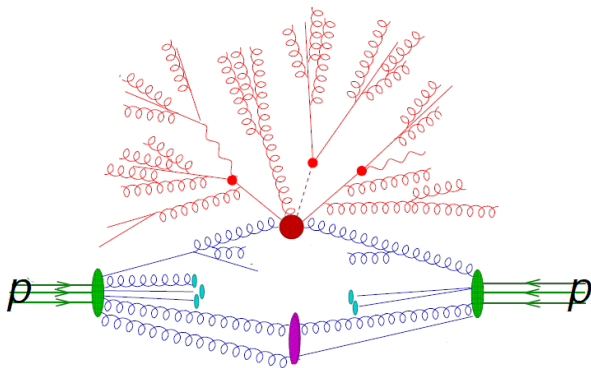


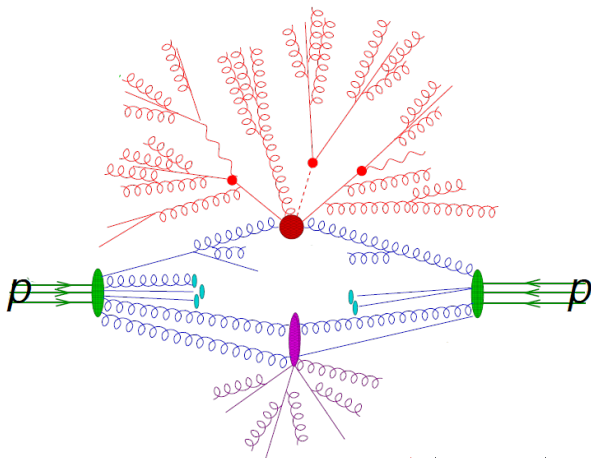
Slika: Primer simulacije dogodka – vsaka barva predstavlja svoj korak





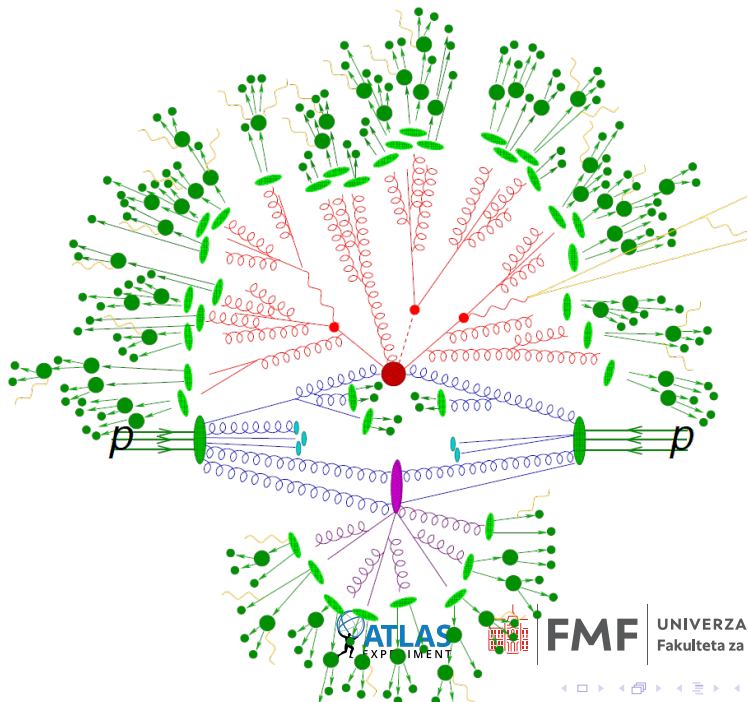






FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko



ATLAS
EXPERIMENT



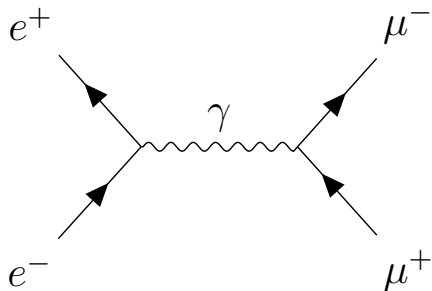
FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

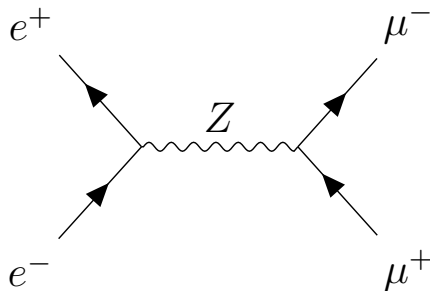


Signal in ozadje

lščemo iglo (signal) v kopici sena (ozadje).

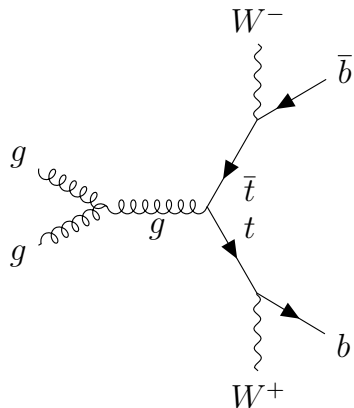


Slika: Primer ozadja za proces $e^+e^- \rightarrow Z \rightarrow \mu^+\mu^-$

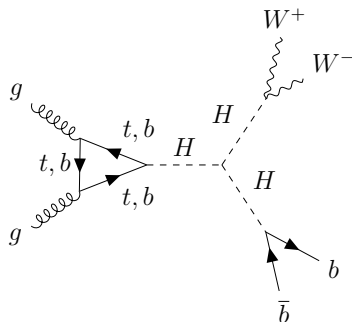


Slika: Primer signala za proces $e^+e^- \rightarrow Z \rightarrow \mu^+\mu^-$

Signal in ozadje



Slika: Primer ozadja za proces $gg \rightarrow HH \rightarrow b\bar{b}WW$



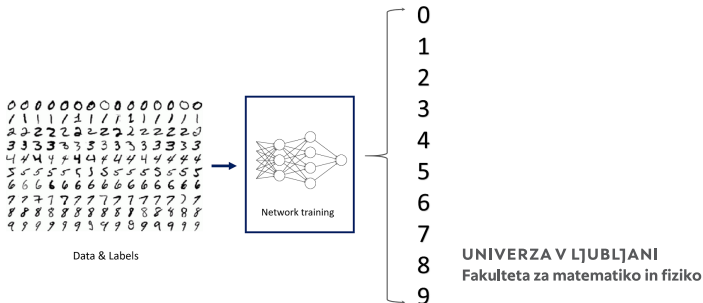
Slika: Primer signala za proces $gg \rightarrow HH \rightarrow b\bar{b}WW$

Strojno učenje

“Strojno učenje je področje, ki se ukvarja z omogočanjem računalnikov, da se učijo, ne da bi jih za to eksplicitno programirali.”

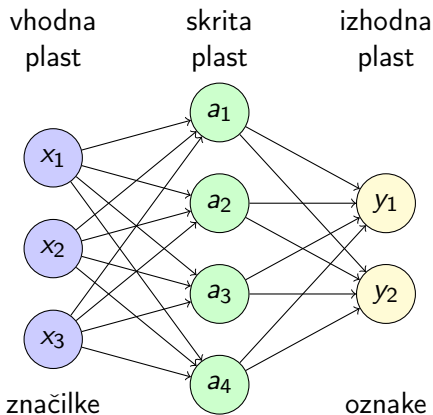
Arthur Samuel, 1959

- Vhodni podatki – *značilke*
- Izhodni podatki – *oznake* ali napovedi
- Primer spodaj: značilke so “slike števil”, oznake pa števila od 1 do 10.

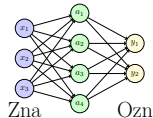


Neuronske mreže

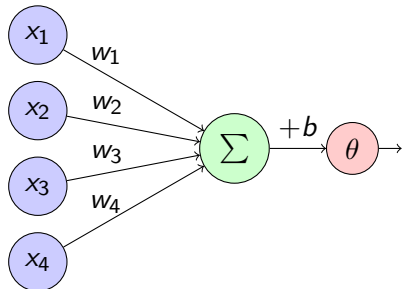
- Povezani sloji nevronov
- Značilke aktivirajo prvi sloj (\mathbf{x}).
- Vsak sloj vpliva na aktivacijo naslednjega sloja.
- Zadnji sloj predstavlja napoved modela – oznake (\mathbf{y}).



Aktivacije



- Aktivacija nevrona odvisna od aktivacij v prejšnjem sloju.
- Linearna kombinacija z utežmi w
- Statistična pristranskost b
- Nelinearna *aktivacijska funkcija* θ
- $a = \theta(x_1 w_1 + \dots + x_n y_n + b)$



x_j : aktivacije prvega sloja;

w_i : uteži;

b : statistična pristranskost;

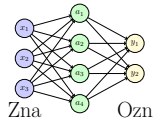
θ : aktivacijska funkcija



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Učenje nevronskih mrež



- Velika količina naborov značilk s podanimi oznakami (\mathbf{y})

L : funkcija izgube;

\mathbf{W} : nabor uteži in pristranskosti;

\mathbf{y} : podane oznake;

$\hat{\mathbf{y}}$: napovedane oznake;

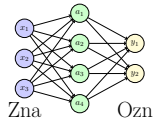
α : korak učenja



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Učenje nevronskih mrež



- Velika količina naborov značilk s podanimi oznakami (\mathbf{y})
- Oznake, ki jih model napove: $\hat{\mathbf{y}}$

L : funkcija izgube;

\mathbf{W} : nabor uteži in pristranskosti;

\mathbf{y} : podane oznake;

$\hat{\mathbf{y}}$: napovedane oznake;

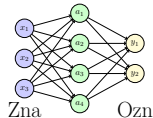
α : korak učenja



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Učenje nevronske mreže



- Velika količina naborov značilk s podanimi oznakami (\mathbf{y})
- Oznake, ki jih model napove: $\hat{\mathbf{y}}$
- Funkcija izgube: $L = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \sum_i (y_i - \hat{y}_i)^2$

L : funkcija izgube;

\mathbf{W} : nabor uteži in pristranskosti;

\mathbf{y} : podane oznake;

$\hat{\mathbf{y}}$: napovedane oznake;

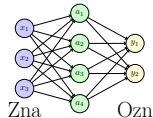
α : korak učenja



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Učenje nevronskih mrež



- Velika količina naborov značilk s podanimi oznakami (\mathbf{y})
- Oznake, ki jih model napove: $\hat{\mathbf{y}}$
- Funkcija izgube: $L = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \sum_i (y_i - \hat{y}_i)^2$
- Funkcija izgube je odvisna od vseh uteži in pristranskosti (\mathbf{W}) v mreži.

L : funkcija izgube;

\mathbf{W} : nabor uteži in pristranskosti;

\mathbf{y} : podane oznake;

$\hat{\mathbf{y}}$: napovedane oznake;

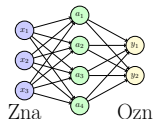
α : korak učenja



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Učenje nevronskih mrež



- Velika količina naborov značilk s podanimi oznakami (\mathbf{y})
- Oznake, ki jih model napove: $\hat{\mathbf{y}}$
- *Funkcija izgube*: $L = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \sum_i (y_i - \hat{y}_i)^2$
- Funkcija izgube je odvisna od vseh uteži in pristranskosti (\mathbf{W}) v mreži.
- Spreminjanje \mathbf{W} , da se minimizira funkcija izgube – učenje.

L : funkcija izgube;

\mathbf{W} : nabor uteži in pristranskosti;

\mathbf{y} : podane oznake;

$\hat{\mathbf{y}}$: napovedane oznake;

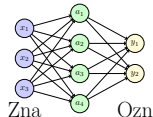
α : korak učenja



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Učenje nevronskih mrež



- Velika količina naborov značilk s podanimi oznakami (\mathbf{y})
- Oznake, ki jih model napove: $\hat{\mathbf{y}}$
- *Funkcija izgube*: $L = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \sum_i (y_i - \hat{y}_i)^2$
- Funkcija izgube je odvisna od vseh uteži in pristranskosti (\mathbf{W}) v mreži.
- Spreminjanje \mathbf{W} , da se minimizira funkcija izgube – učenje.
- $\mathbf{W}_{\text{nove}} = \mathbf{W}_{\text{stare}} + \alpha \nabla L(\mathbf{W})$

L : funkcija izgube;

\mathbf{W} : nabor uteži in pristranskosti;

\mathbf{y} : podane oznake;

$\hat{\mathbf{y}}$: napovedane oznake;

α : korak učenja



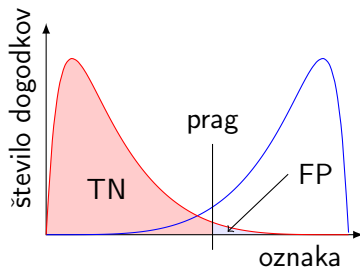
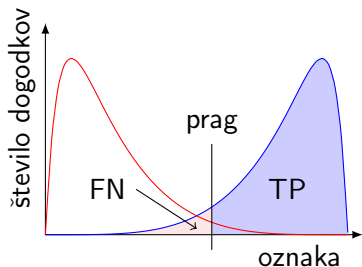
FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

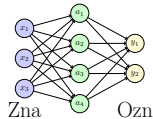
Mere za uspešnost modelov

Skozi natrenirano mrežo pošljemo veliko naborov značilnik z znanimi oznakami. Ločimo štiri vrste dogodkov:

- *pravilno pozitivni* (TP): model je pravilno napovedal prisotnost bozona,
- *lažno pozitivni* (FP): model je napovedal prisotnost bozona, v resnici pa ga ni bilo,
- *pravilno negativni* (TN): model je pravilno napovedal odsotnost bozona,
- *lažno negativni* (FN): model je napovedal odsotnost bozona, v resnici pa je bil.



Mere za uspešnost modelov



- Delež pravilno pozitivnih primerov (občutljivost)

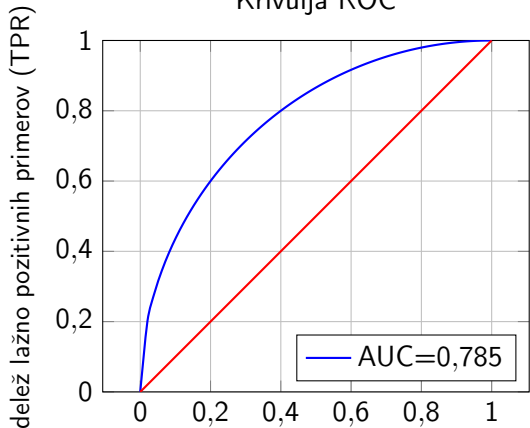
$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- Delež lažno pozitivnih primerov

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

- Parametrična krivulja ROC
- Ploščina pod krivuljo (AUC)

Krivulja ROC







delež pravilno pozitivnih primerov (FPR)



FMF

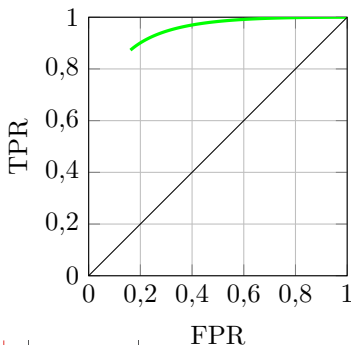
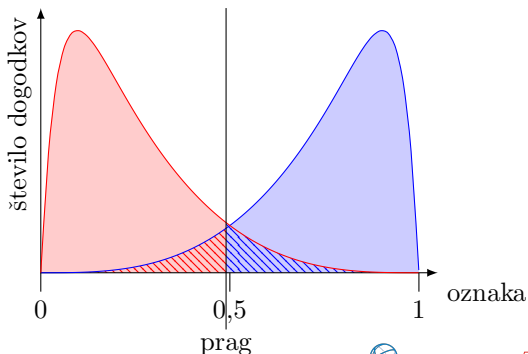
UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

-  FP: lažno pozitivni primeri
-  FN: lažno negativni primeri
-  TN: pravilno negativni primeri
-  TP: pravilno pozitivni primeri

prag = 0,490

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{light blue}}{\text{light blue} + \text{red hatching}} = 0,865$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} = \frac{\text{blue hatching}}{\text{blue hatching} + \text{light red}} = 0,153$$



FMF

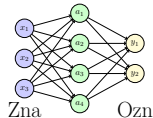
FMR
UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko



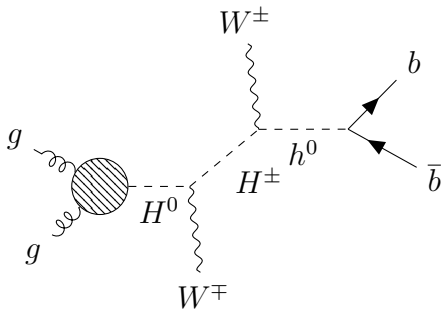
FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

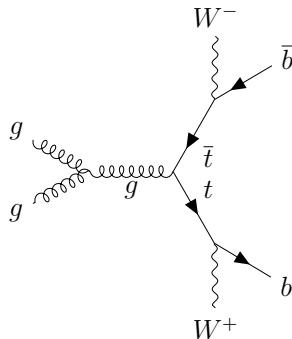
Primer uporabe: globoke mreže



Nastanek hipotetičnih težkih Higgsovih bozonov H^0, H^+, H^- .



Slika: Proces signala.



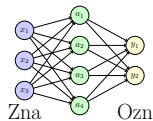
Slika: Proces ozadja.



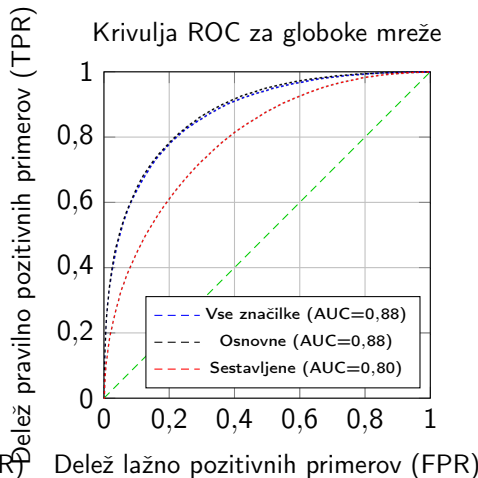
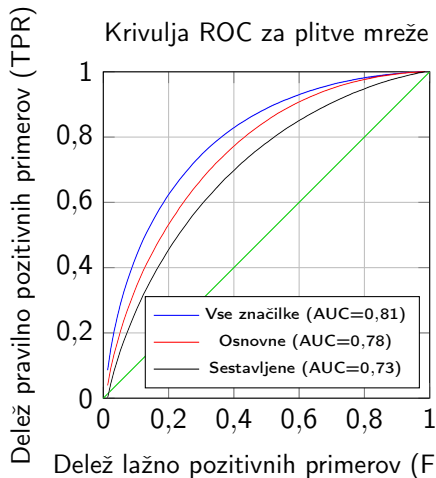
FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

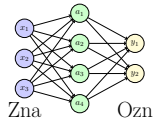
Primer uporabe: globoke mreže



- Uporabljene značilke: 21 osnovnih in 7 sestavljenih.
- Dve vrsti mrež: plitve in globoke.



Avtoenkoderji



Poseben primer nevronske mreže. Učenje na ozadju in zaznavanje odstopenj.

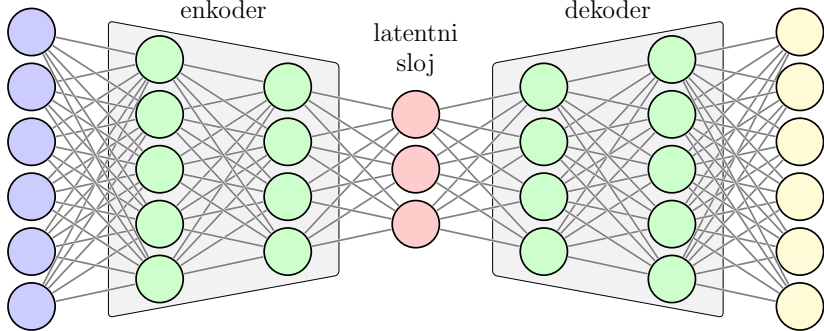
vhodni sloj

enkoder

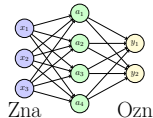
latentni sloj

dekoder

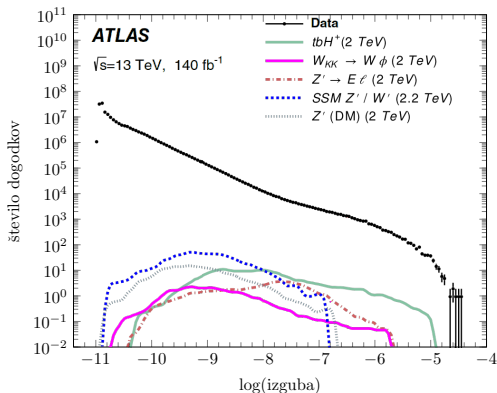
izhodni sloj

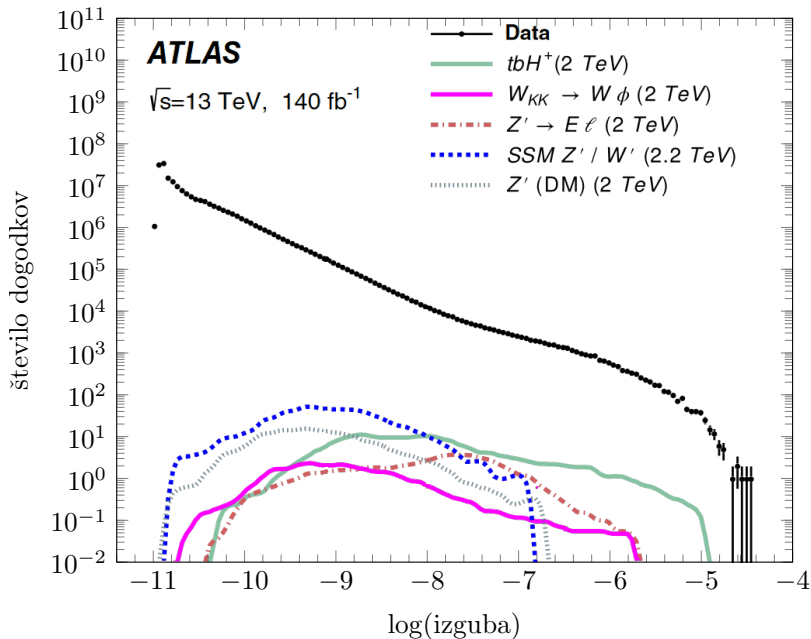


Primer uporabe avtoenkoderjev



- 1278 značilk \mathbf{x} , 1278 oznak \mathbf{y}
- Latentni sloj dimenzije 200
- Funkcija izgube
 $L = \text{izguba} = \|\mathbf{x} - \mathbf{y}\|^2$
- Treniranje na simulacijah Standardnega modela
- Preizkus na pravih meritvah.
- Premalo meritev za statistično dovolj gotovo potrditev.





Zaključek

- Obetavno področje.
- Deluje v principu, ni še veliko uporabnih rezultatov.
- Zelo hiter razvoj tehnologij in metod.
- Potrebni več meritev ali višje energije.
- Na voljo vedno bolj izpopolnjena orodja.
- S tem se ukvarjajo tudi na IJS (F9).



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko



Hvala za pozornost!



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko



Dodatek: luminoznost

Pričakovano število dogodkov (trkov) N za določen proces:

$$N = \sigma \int \mathcal{L} dt = \sigma L,$$

kjer je \mathcal{L} luminoznost, L pa integrirana luminoznost.

$$\mathcal{L} = \frac{\mathcal{N}^2 n_b f}{A},$$

kjer je \mathcal{N} število protonov v gruči, n_b število gruč, A pa efektiven presek žarka na mestu trka.



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko

Dodatek: nevronske mreže bolj matematično

Nevronska mreža je funkcija $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, ki pretvori značilke $\mathbf{x} \in \mathbb{R}^n$ v oznake $\mathbf{y} \in \mathbb{R}^m$. Aktivacije i -tega sloja $\mathbf{a}^{(i)}$ se lahko zapišejo kot aktivacije $i - 1$ -tega sloja $\mathbf{a}^{(i-1)}$ kot

$$\mathbf{a}^{(i)} = \theta(W\mathbf{a}^{(i-1)} + \mathbf{b}) = \sigma_n(\mathbf{a}^{(i-1)}).$$

Vektor \mathbf{b} so pristranskosti, matrika W so uteži, θ pa je aktivacijska funkcija. Prehod med slojema označimo s σ_j . Celotno mrežo lahko zapišemo kot

$$\mathbf{y} = f(\mathbf{x}) = (\sigma_j \circ \sigma_{j-1} \circ \cdots \circ \sigma_2 \circ \sigma_1)(\mathbf{x}),$$

kjer je j število slojev.



FMF

UNIVERZA V LJUBLJANI
Fakulteta za matematiko in fiziko