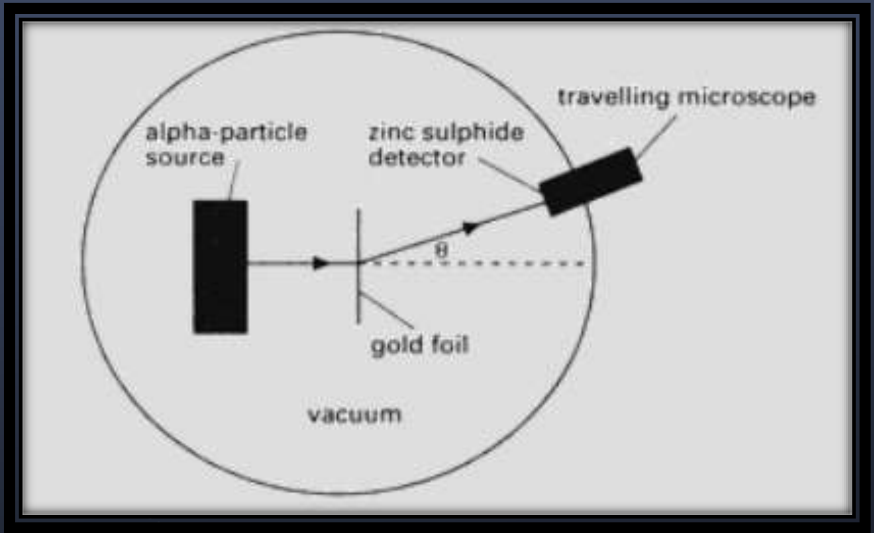
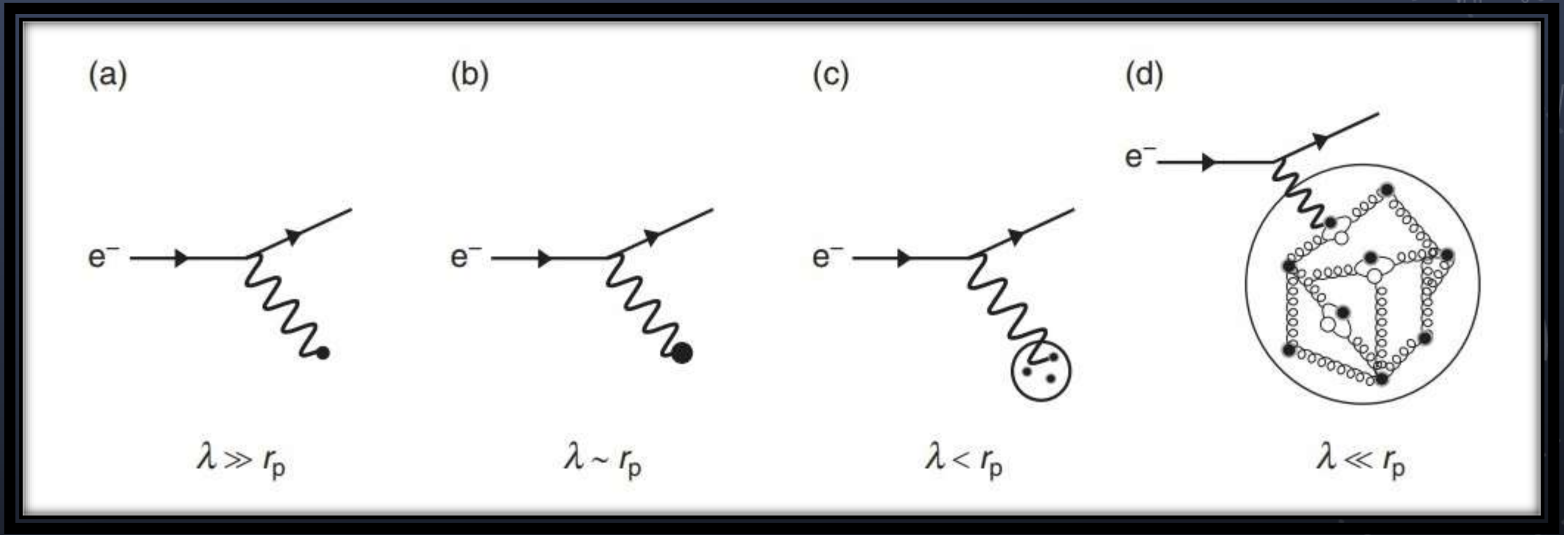
Razotkrivanje strukture hadrona gausijanskim procesom

Andrija Radočaj

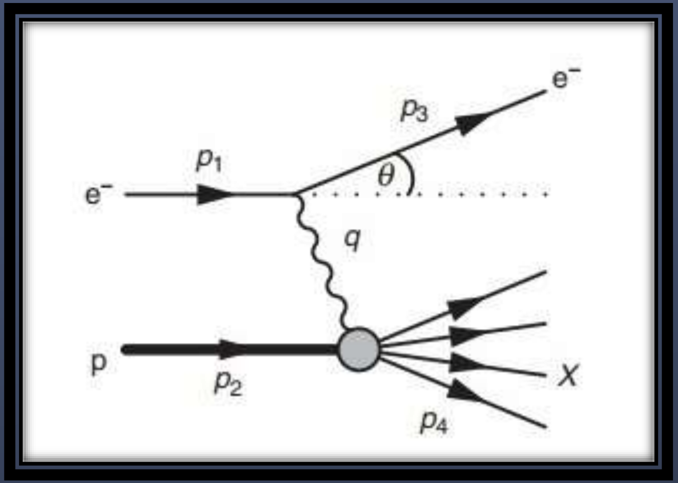
Procesi raspršenja

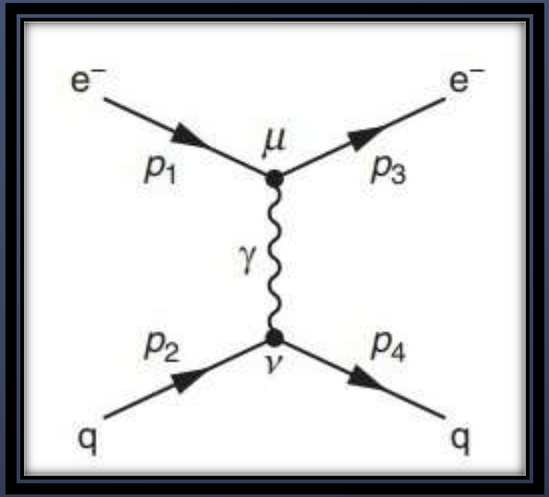
- E. Rutherford 1909. godina – struktura atoma
- SLAC 1960.-ih godina – struktura protona





Duboko neelastično raspršenje





- Kinematičke varijable x, Q^2, y, υ
- Partonske distribucijske funkcije $q^p(x)$ – raspodjela impulsa

$$\frac{d^2\sigma}{dx dQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[(1-y)+\frac{y^2}{2}\right]\sum_i Q_i^2 q_i^p(x)$$

Strojno učenje

MODEL + PARAMETRI

Učenje pod nadzorom

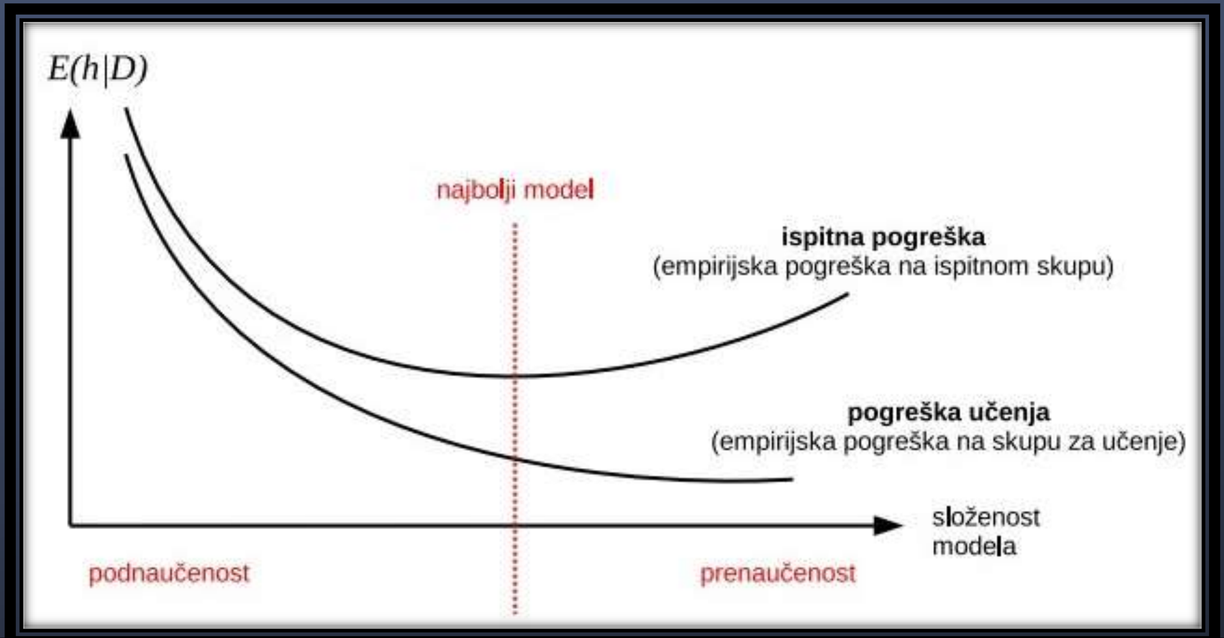
- Skup za učenje $\{(x_i, y_i), i = 1, \dots, n\}$
- Regresija, klasifikacija

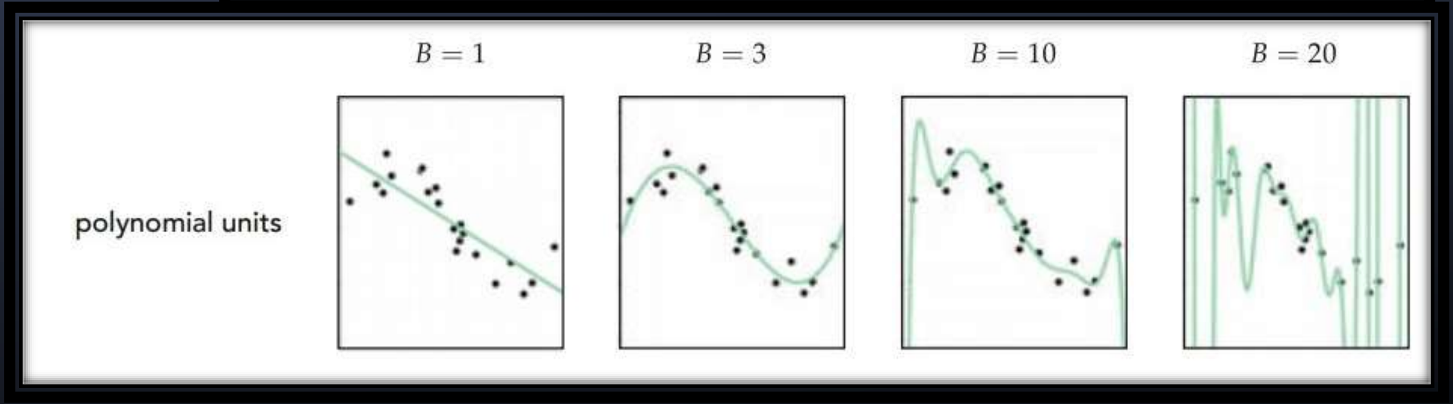
Učenje bez nadzora

- Skup za učenje neoznačen

Optimizacija modela

- Minimizacija funkcije gubitka/pogreške





Gradijentni spust (1)

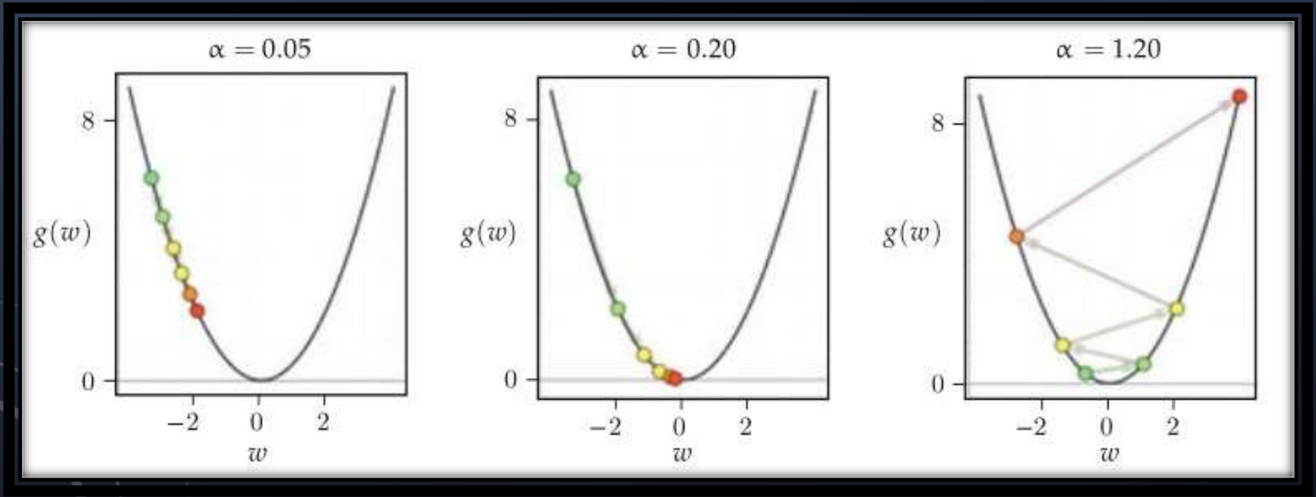
Ulaz: skup za učenje (X, Y), diferencijabilna funkcija $\mathcal{L}(X, Y, \vec{\boldsymbol{w}})$, stopa učenja $\eta^{(i)}$, početna parametarska točka $\vec{\boldsymbol{w}}^{(0)} = \left(w_1^{(0)}, \dots, w_n^{(0)}\right)$

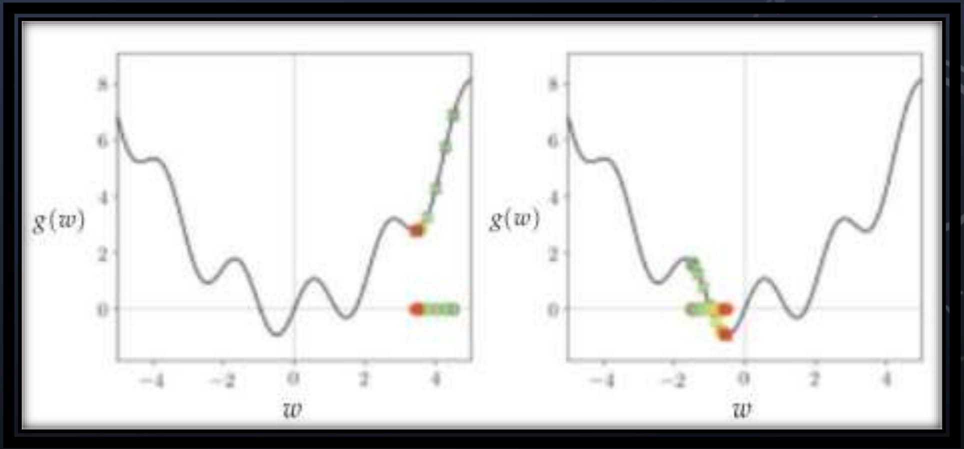
1. $i = 0$
2. **while** uvjet zaustavljanja nije ispunjen **do**:
3. Računanje $\nabla_{\vec{\boldsymbol{w}}} \mathcal{L}(X, Y, \vec{\boldsymbol{w}}^{(i)})$
4. Nadogradnja $\vec{\boldsymbol{w}}^{(i+1)} = \vec{\boldsymbol{w}}^{(i)} - \eta^{(i)} \nabla_{\vec{\boldsymbol{w}}} \mathcal{L}(X, Y, \vec{\boldsymbol{w}}^{(i)})$
5. $i \leftarrow i + 1$
6. **end while**

Gradijentni spust (2)

Mogući uvjeti zaustavljanja **while** petlje:

- $\left\|\nabla_{\vec{\boldsymbol{w}}}\mathcal{L}(X,Y,\vec{\boldsymbol{w}}^{(i)})\right\|_2 \approx 0$
- $\left\|\vec{\boldsymbol{w}}^{(i+1)} - \vec{\boldsymbol{w}}^{(i)}\right\|_2 \approx 0$
- $\left|\mathcal{L}\left(X,Y,\vec{\boldsymbol{w}}^{(i+1)}\right) - \mathcal{L}(X,Y,\vec{\boldsymbol{w}}^{(i)})\right| \approx 0$





Gausijanski proces

- Skup nasumičnih varijabli $\vec{\boldsymbol{f}}_D = (f_1, \dots, f_D) = \left(f(\vec{\boldsymbol{x}}_1), \dots, f(\vec{\boldsymbol{x}}_D)\right)$, svaki njihov konačan broj čini multivarijantnu Gaussovu raspodjelu

- Multivarijantna Gaussova raspodjela za D-dim nasumični vektor:

$$\vec{\boldsymbol{f}}_D \sim \mathcal{N}_D(\vec{\boldsymbol{\mu}}, \Sigma)$$

$$\vec{\boldsymbol{\mu}} = E\left[\vec{\boldsymbol{f}}\right] = \left(E(f_1), \dots, E(f_D)\right)^T = \left(E[f(\vec{\boldsymbol{x}}_1)], \dots, E[f(\vec{\boldsymbol{x}}_D)]\right)^T \equiv \left(\mu(\vec{\boldsymbol{x}}_1), \dots, \mu(\vec{\boldsymbol{x}}_D)\right)^T$$

$$\Sigma_{i,j} = E\left[(f(\vec{\boldsymbol{x}}_i) - \mu_i)(f(\vec{\boldsymbol{x}}_j) - \mu_j)\right] = Cov\left[f(\vec{\boldsymbol{x}}_i), f(\vec{\boldsymbol{x}}_j)\right] = k(\vec{\boldsymbol{x}}_i, \vec{\boldsymbol{x}}_j)$$

- Gausijanski proces: $f(\vec{\boldsymbol{x}}) \sim \mathcal{GP}(\mu(\vec{\boldsymbol{x}}), k(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}'))$

$$exp.quad(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}'; l) = \exp\left(-\frac{1}{2l^2}|\vec{\boldsymbol{x}} - \vec{\boldsymbol{x}}'|^2\right)$$

Regresija gausijanskih procesa (1)

$$\{(\vec{\boldsymbol{x}}_i, y_i) | i = 1, \ldots, n\}, \quad y_i = f(\vec{\boldsymbol{x}}_i) + \varepsilon$$

$$cov(y_i, y_j) = k(\vec{\boldsymbol{x}}_i, \vec{\boldsymbol{x}}_j) + \sigma_n^2 \delta_{i,j} \rightarrow \; cov(\vec{\boldsymbol{y}}) = K(X, X) + \sigma_n^2 I$$

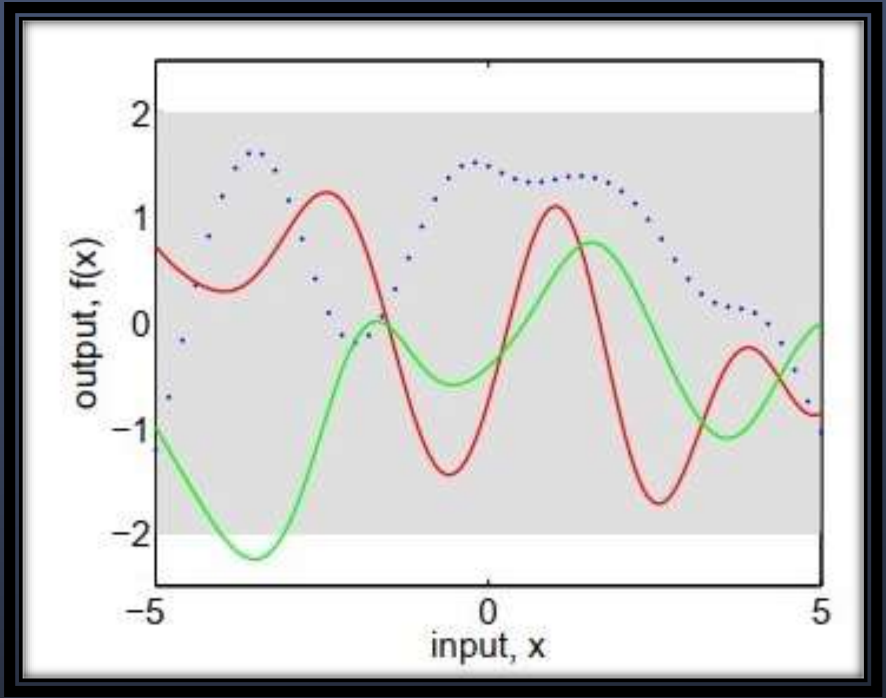
$$\begin{bmatrix} \vec{\boldsymbol{y}} \\ \vec{\boldsymbol{f}}_* \end{bmatrix} \sim \mathcal{N}\left(\vec{\mathbf{0}}, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right)$$

$$\vec{\boldsymbol{f}}_* | X, \vec{\boldsymbol{y}}, X_* \sim \mathcal{N}\left(\overline{\vec{\boldsymbol{f}}_*}, cov(\vec{\boldsymbol{f}}_*)\right)$$

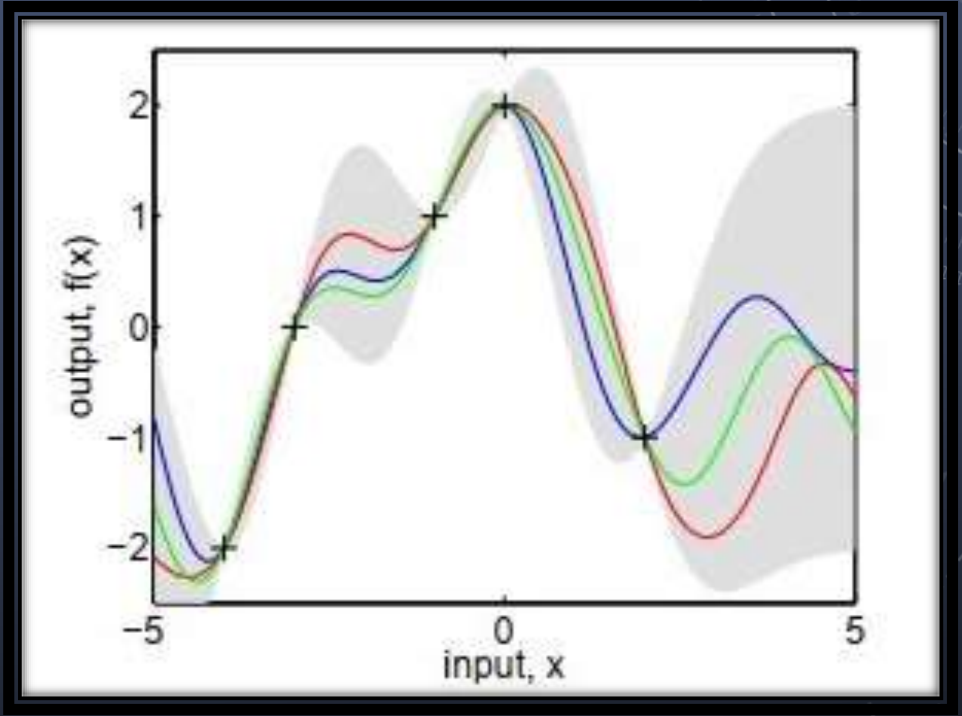
$$\overline{\vec{\boldsymbol{f}}_*} = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} \vec{\boldsymbol{y}}$$

$$cov(\vec{\boldsymbol{f}}_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X_*)$$

Regresija gausijanskih procesa (2)



$\overrightarrow{f_*} \sim \mathcal{N}(\vec{\mathbf{0}}, exp.quad(x, x'; l = 1))$



$\overrightarrow{f_*} \sim \mathcal{N}(\vec{\mathbf{0}}, exp.quad(x, x'; l = 1))$

+ mjerenja bez greški

Granična vjerojatnost

Pretpostavljeni model: $\vec{\boldsymbol{f}}|X \sim \mathcal{N}_n(\vec{\mathbf{0}}, K_{\vec{\boldsymbol{w}}})$

$$y_i = f(x_i) + \varepsilon \quad \rightarrow \quad \vec{\boldsymbol{y}}|\vec{\boldsymbol{f}}, X \sim \mathcal{N}_n(\vec{\boldsymbol{f}}, \sigma_n^2 I)$$

$$p(\vec{\boldsymbol{y}}|X) = \int p(\vec{\boldsymbol{y}}|\vec{\boldsymbol{f}}, X)\, p(\vec{\boldsymbol{f}}|X)\, d\vec{\boldsymbol{f}}$$

$$ln[p(\vec{\boldsymbol{y}}|X)](\vec{\boldsymbol{w}}) = -\frac{1}{2}\vec{\boldsymbol{y}}^T(K_{\vec{\boldsymbol{w}}} + \sigma_n^2 I)^{-1}\vec{\boldsymbol{y}} - \frac{1}{2} ln|K_{\vec{\boldsymbol{w}}} + \sigma_n^2 I| - \frac{n}{2} ln(2\pi)$$

$\mathcal{L}(\vec{\boldsymbol{w}}) = (-1) \cdot ln[p(\vec{\boldsymbol{y}}|X)](\vec{\boldsymbol{w}}) \rightarrow$ funkcija gubitka

Pojednostavljenje problema ekstrakcije

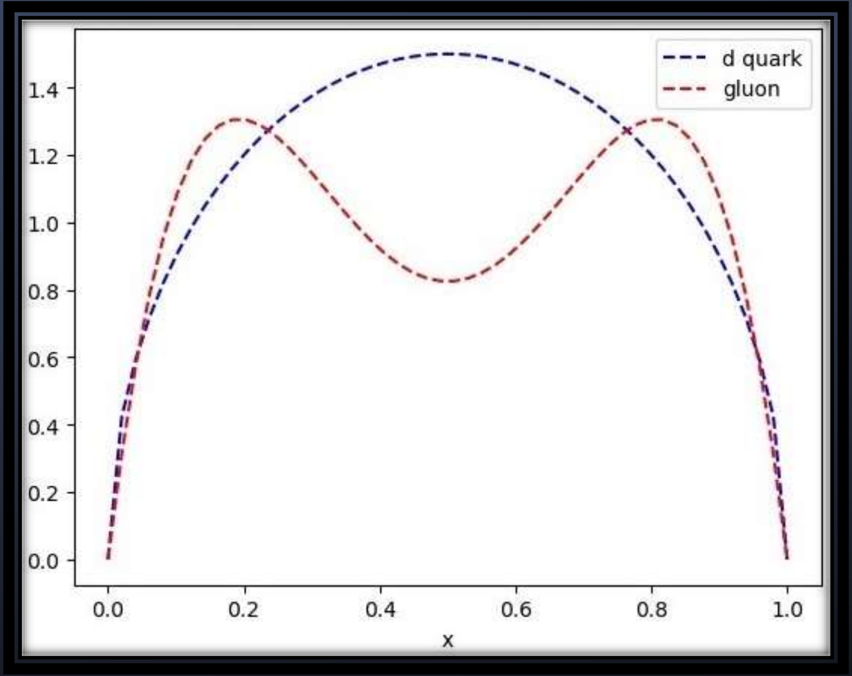
- $$\frac{d\sigma_1}{dx} = h_1(x)[0] \cdot dquark_{truth}(x) + h_1(x)[1] \cdot gluon_{truth}(x)$$
- $$\frac{d\sigma_2}{dx} = 1 + h_2(x)[0] \cdot dquark_{truth}(x) + h_2(x)[1] \cdot gluon_{truth}(x)$$

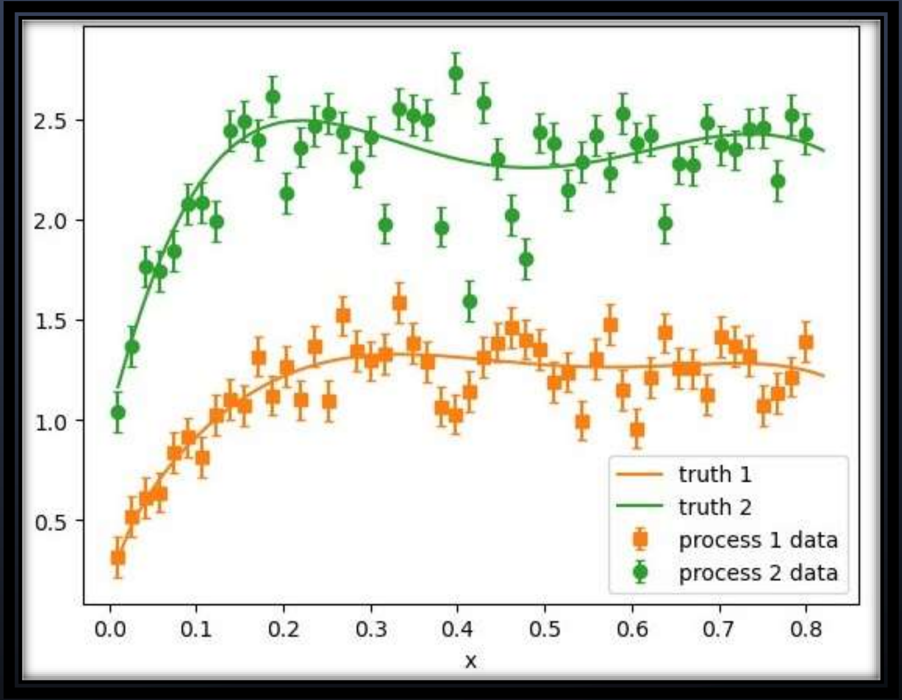
$$h_1(x) = \left[\frac{1}{1+x}, \frac{x}{1+x}\right] \quad , \quad h_2(x) = \left[\frac{x}{1+x^2}, \frac{1}{1+x^2}\right]$$

Simulacija eksperimentalnih podataka

$$dquark_{truth}(x) = 3\sqrt{x \cdot (1 - x)}$$

$$gluon_{truth}(x) = 6x \cdot (1 - x) \cdot \left(1 + 0.3 \cdot C_2^{(1.5)}(2x - 1)\right)$$





PDF - nezavisni gausijanski procesi (1)

- dquark $\rightarrow exp.quad(x, x'; l_1)$
- gluon $\rightarrow exp.quad(x, x'; l_2)$

$$\frac{d\sigma_1}{dx} = h_1(x)[0] \cdot dquark_{truth}(x) + h_1(x)[1] \cdot gluon_{truth}(x)$$

$$cov_1(x, x') = h_1(x)[0]h_1(x')[0] \cdot exp.quad(x, x'; l_1) + h_1(x)[1]h_1(x')[1] \cdot exp.quad(x, x'; l_2)$$

$$\frac{d\sigma_2}{dx} - 1 = h_2(x)[0] \cdot dquark_{truth}(x) + h_2(x)[1] \cdot gluon_{truth}(x)$$

$$cov_2(x, x') = h_2(x)[0]h_2(x')[0] \cdot exp.quad(x, x'; l_1) + h_2(x)[1]h_2(x')[1] \cdot exp.quad(x, x'; l_2)$$

PDF - nezavisni gausijanski procesi (2)

$$\begin{cases} \dfrac{d\sigma_1}{dx} = h_1(x)[0] \cdot dquark_{truth}(x) + h_1(x)[1] \cdot gluon_{truth}(x) \\ \dfrac{d\sigma_2}{dx} - 1 = h_2(x)[0] \cdot dquark_{truth}(x) + h_2(x)[1] \cdot gluon_{truth}(x) \end{cases}$$

↳

$$cov_{12}(x,x') = h_1(x)[0]h_2(x')[0] \cdot exp.quad(x,x';l_1) + h_1(x)[1]h_2(x')[1] \cdot exp.quad(x,x';l_2)$$

↳

$$cov_{21}(x,x') = h_2(x)[0]h_1(x')[0] \cdot exp.quad(x,x';l_1) + h_2(x)[1]h_1(x')[1] \cdot exp.quad(x,x';l_2)$$

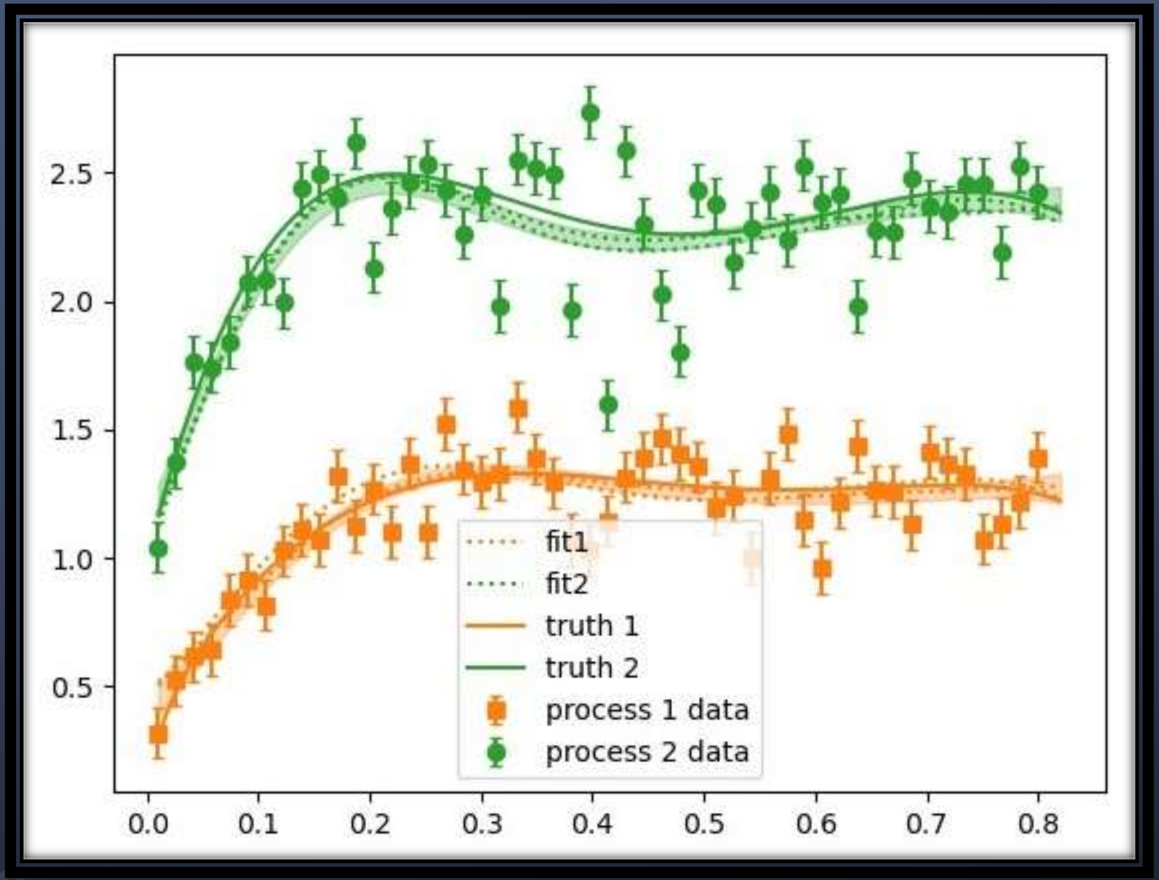
Matrica kovarijance

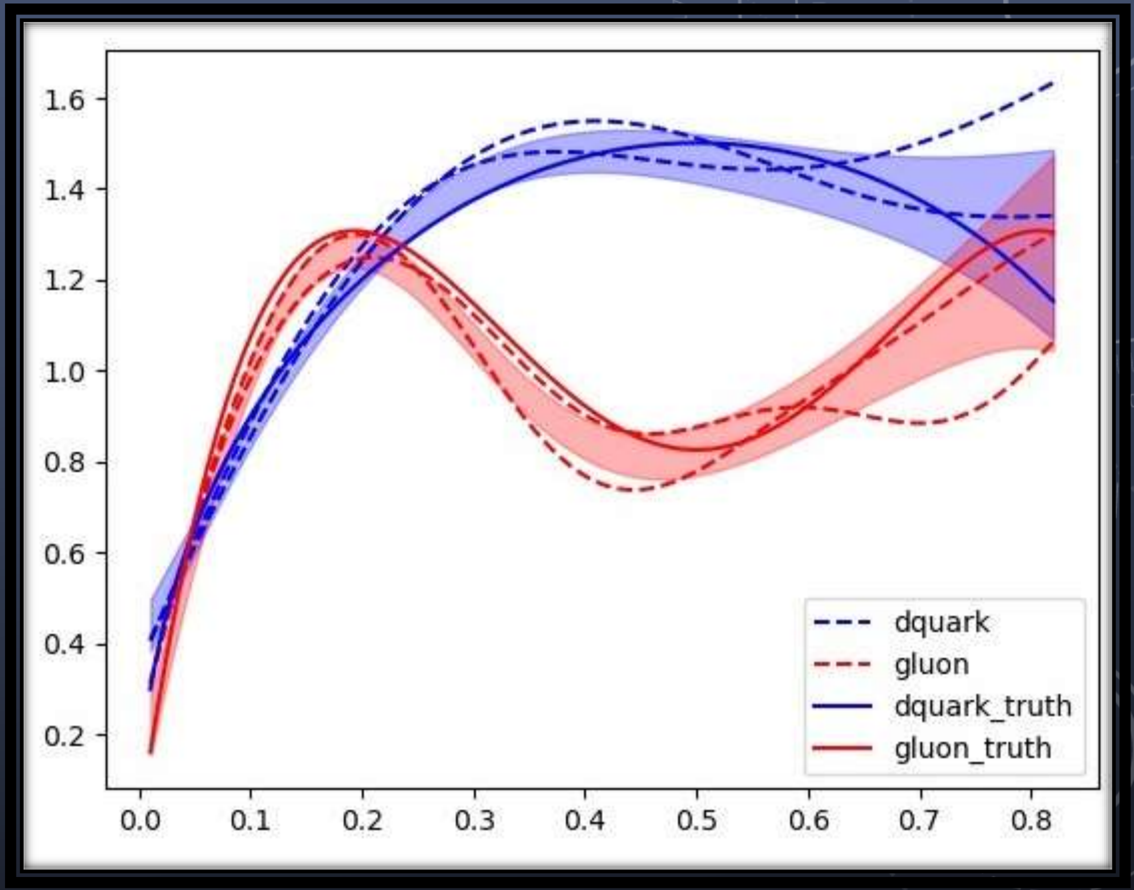
- $\vec{\boldsymbol{y}} = [y_1^1, \dots, y_n^1, y_1^2 - 1, \dots, y_n^2 - 1]^T$,
  y_i^p - srednja vrijednost mjerenja za $\left.\frac{d\sigma_p}{dx}\right|_{x_i}$;

- $$K = \begin{bmatrix} cov_1(x_1,x_1) & \cdots & cov_1(x_1,x_n) & cov_{12}(x_1,x_1) & \cdots & cov_{12}(x_1,x_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ cov_1(x_n,x_1) & \cdots & cov_1(x_n,x_n) & cov_{12}(x_n,x_1) & \dots & cov_{12}(x_n,x_n) \\ cov_{21}(x_1,x_1) & \cdots & cov_{21}(x_1,x_n) & cov_2(x_1,x_1) & \cdots & cov_2(x_1,x_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ cov_{21}(x_n,x_1) & \cdots & cov_{21}(x_n,x_n) & cov_2(x_n,x_1) & \cdots & cov_2(x_n,x_n) \end{bmatrix}$$

$\vec{\boldsymbol{y}}$; $K = K\left(\begin{bmatrix} l_1 \\ l_2 \end{bmatrix}\right) = K(\vec{\boldsymbol{w}})$ → podatci za funkciju gubitka → optimizacija $\vec{\boldsymbol{w}}$

Rezultati





Prilagodba nakon optimizacije ➡ Ekstrakcija rješavanjem sustava

Zaključak

- Metoda ekstrakcije pdf-a iz podataka dif. udarnog presjeka:

1. Pojednostavljenje problema
2. Simulacija podataka dif. udarnog presjeka zadavanjem pdf-a (Python) → cilj replikacija zadanih pdf-a
3. Modeliranje pdf-a kao gausijanskih procesa (lsqfitgp paket)
4. Optimizacija parametara matrice kovarijance na simulirane podatke (lsqfitgp paket)
5. Prilagodba na simulirane podatke (lsqfitgp paket)
6. Ekstrakcija pdf-a (lsqfitgp paket) → obećavajući rezultati

- Preinake i nadogradnje → primjena u istraživanju

Literatura

[1] A. Isaacs, *Oxford Dictionary of Physics*. Oxford University Press, 2000.

[2] M. Thomson, *Modern Particle Physics*. Cambridge University Press, 2013.

[3] J. Šnajder, Bilješke s predavanja – strojno učenje. Dostupno na: https://www.fer.unizg. hr/predmet/struce1/materijali#%23!p_rep_142310!_-210848.

[4] J. Watt, R. Borhani i A. K. Katsaggelos, *Machine Learning Refined: Foundations, Algorithms and Applications*. Cambridge University Press, 2020.

[5] C. E. Rasmussen i C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[6] lsqfitgp Manual. Dostupno na: https://gattocrucco.github.io/lsqfitgp/docs/.