

# Kvantno masinsko učenje

Milan Bojic

Jun 2022

# Sadržaj

<b>1</b>	<b>Kvantno racunarstvo</b>	<b>3</b>
1.1	Osnovni pojmovi . . . . .	3
1.2	Kvantno racunarstvo . . . . .	5
1.3	Kvantna inforamcija . . . . .	5
1.3.1	Priprema podataka . . . . .	6
<b>2</b>	<b>Linearne algebra za kvantno masinsko učenje</b>	<b>6</b>
<b>3</b>	<b>Kvantna teorija kompleksnosti</b>	<b>7</b>
<b>4</b>	<b>Kvantno masinsko učenje</b>	<b>7</b>
4.1	Quantum support vector machine . . . . .	8
4.1.1	Klasican algoritam . . . . .	8
4.1.2	Kvantni algoritam . . . . .	9
4.2	Quantum principal component analysis . . . . .	11

# 1 Kvantno racunarstvo

Pre nego sto se pocne pricati o Kvantnom masinskom ucenju, treba objasniti neki osnovni pojmovi da bi lakse razumeli ostatak rada

## 1.1 Osnovni pojmovi

Potrebni pojmovi su:

- Kubit (eng. Qubit)
- Kvantna kola (eng. Quantum Gates)
- Kvantna uvezanost (eng. Quantum entanglement)
- Kvantan memorija, Kvantni registri

### Kubit

Kubit (eng. Qubit ) je najmanja jedinica informacije u kvantnom računarstvu, slično bit-u u klasičnom računarstvu. Razlika od bita jeste u tome što kubit pored stanja 1 i 0, može da se nalazi i u superpoziciji između oba. Oni se mogu predstaviti formulom (koristeci "bra-ket" notaciju):

$$|\gamma\rangle = \alpha |0\rangle + \beta |1\rangle$$

Ovde su  $|0\rangle$  i  $|1\rangle$  zapravo stanja kao i kod klasičnog bita, a  $\alpha$  i  $\beta$  su kompleksni brojevi koji predstavljaju amplitude zadatih stanja i za njih važi:

$$|\alpha|^2 + |\beta|^2 = 1$$

Posto stanje kubita ima dva stepena slobode sto dovodi do toga da amplitude se mogu zapisati kao:

$$\alpha = \cos \frac{\Theta}{2}$$
$$\beta = e^{i\phi} \sin \frac{\Theta}{2}$$

Takodje mozemo da vidimo da je  $|\alpha|^2$  verovatnoca da se kubit nalazi u stanju 0, isto vazi i za  $|\beta|^2$  i 1. Saznanje o tomo u kom stanju se nalazi kubit ce se dobiti merenjem kubita, tade ce da kubit izadje iz superpozicije i "pasce" u

stanje 1 ili stanje 0. U tom slučaju kubit će imati ponašanje kao i običan bit, ali ovako gubimo predjasnje kvantno stanje kubita. U fizičkom svetu kubit se može predstaviti kao polarizovani foton, pre čemu se dva stanja se uzimaju kao vertikalna i horizontalna polarizacija.

## Kvantna kapija

Kvantna kapije (eng. Quantum Gates) su logički predstavljaju matricama i oni rade na određenom broju kubita. Matrice su unitarne sa oblikom  $2^n \times 2^n$ , gde je  $n$  broj qubita na kojim radimo. Neke od poznatih kola su: Hadamardovo kolo (stavlja kubit u superpoziciju), bit flip kolo (zamenjuje amplitude na kubit), ali nas najviše zanima rotaciono kolo:

$$R = \begin{bmatrix} \cos \Theta & -\sin \Theta \\ \sin \Theta & \cos \Theta \end{bmatrix}$$

Ovo kolo rotira kubite u prostoru, odnosno menja njihove amplitude za  $\Theta$  radiana.

## Kvantna uvezanost

Kvantna uvezanost (eng. Quantum entanglement) je fizički pojam gde su dva, ili više, kubita povezana tako da zajedno prave novo kvantno stanje. U čistim stanjima oni su matematički zapravo proizvodi tenzora amplitude:

$$|\gamma\rangle \otimes |\delta\rangle = \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle$$

I ovako napisano kvantno stanje se može razdvojiti na dva kubita. Ali postoje i kvantna stanja koja se ne mogu razdvojiti npr.

$$\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$$

Zanimljiva stvar kod uvazanih kubita jeste u tome što dele informacije. Ako bi jedan kubit iz para odneli u neko veoma daleko mesto (na primer druga galaksija), i tamo bi ga izmerili mi bi smo dobili 0 ili 1, međutim drugi kubit bi takođe upao u određeno stanje i to u istom trenutni kad smo izmerili prvi daleki kubit. Ovo je zapravo gde se nalazi glavna različitost između klasičnog i kvantnog računarstva, ova pojava ne postoji u klasičnom računarstvu i ne može se "lako" simulirati.

## Kvantni registri

Kvantni registri se sastoje od kvantnog stanja od  $m$  uvezanih kubita i može da se predstavljati do  $2^m$  vrednosti stanja istovremeno. Kvantna memorija su uređaji koji čuvaju kvantna stanja fotona, bez da uništavaju kvantnu informaciju koja se nalazi u fotonu. Ovakva memorija zahteva koherentni sistem materije, jer bi u suprotnom kvantna informacija unitar uređaja bila izgubljena zbog nekoherentnosti.

## 1.2 Kvantno računarstvo

Kvantno računarstvo je vrsta računarstva gde se koriste kolekcije fizickih osobina kvantne mehanike kao što su superpozicija i kvantna uvezanost, tako da se izvrši neka kalkulacija. Uređaja koji izvršava kvantne kalkulacije zovu se **kvantni računari**. Kvantni računari se sastoje od kvantnih kola i elementarnih kvantnih kapija koje služe za prenos i manipulisanje kvantnih informacija. [6]

Jedna od glavnih primena Kvantnih računara jeste simulacija fizickih sistema, bilo oni kvantne ili klasične prirode.

## 1.3 Kvantna informacija

Kvantna informacija je informacija o stanju kvantnog sistema. O njihovim svojstvima bavi se **kvantna teorija informacije**. Takođe, kvantna informacija mogu izmeriti na isti način kao i klasična informacija koristeći se Šenonovom metodom. Odnosno, postoji jedinstveno merilo to jest funkcija nad kvantnim stanjem, koje je funkcija verovatnoće, kontinuiteta i sumiranja.[8] Ova funkcija se zove **von Neumann entropy** i za neki ulazni kubit  $\rho$  postoji ekvivalent u **Shannon entropy**  $H$  za neku slučajnu promenljivu  $X$

$$S(\rho) = H(X)$$

Jos jedna od merila za kvantno stanje jeste merenje "istinitosti" (eng. Fidelity) između dva kvantna stanja  $|\phi\rangle$  i  $|\psi\rangle$ . Neka je  $F$  funkcija koja meri osobinu, ona meri verovatnoću da merenjem stanja  $|\phi\rangle$  dobijemo stanje  $|\psi\rangle$ . Izlaz funkcije je između 0 i 1, gde ako je izlaz 0 onda su dva stanja ortogonalna jedna od drugog, a ako je izlaz 1 onda su dva stanja jednaka.[8]

## Odnost kvantne i klasicne teorije informacije

Kvantna i klasicna informacija se u dosta stvari razlikuju. Dok se klasicna informacija prolazi kroz sisteme sa dobro definisanim stanjima, moze se kopirati i pri procesu merenja se nemanja, Kvantna informacija je enkodovana u kvantnim sistemima, ne moze se kopirati i pri procesu merenja ona se menja. Takodje kvantna informacija ima neke osobine koje se ne mogu iskazati u klasicnoj informaciji, kao sto su superpozicija i kvantna uvezanost [4] Kvantna teorija informacije se bavi:

1. Prenosjenju klasicnih informacija preko kvantnih kanala
2. Prenosjenju kvantnih informacija preko kvantnih kanala
3. Efekat kvantne uvezanosti na prenosjenje informacija

### 1.3.1 Priprema podataka

Za obradu podataka treba nam kvantni RAM (QRAM), koji nam dozvoljava paralelan pristup kvantnim podacima. Neka imamo kompleksan vektor  $\vec{v}$  sa  $N = 2^n$  dimenzija, gde su njegove komponente oblika

$$v_j = |v'_j|e^{i\Phi_j}$$

Ako imamo parove  $|v'_j|, \Phi_j$  cuvamo kao float brojeve u QRAM-u, onda mozemo da konstruisemo  $\log_2 N$  kubit kvantno stanje  $|v\rangle = |\vec{v}|^{-\frac{1}{2}} \vec{v}$  u  $O(\log N)$  koraka

Kada smo kreirali kompresovane kvantne vektore od ulaznih vektora, mozemo da vrsimo transformacije koristeći kvantne algoritme, za dalje koriscenje podataka za masinsko ucenje. Ovaj proces zove se **postprocessing** i u opstem obliku njemu je potrebno  $O(poly(\log N))$  koraka. [3]

## 2 Linearne algebra za kvantno masinsko ucenje

Da bi videli kako kvantni racunari poboljsavaju masinsko ucenje, treba da se vidi kako kvantni racunari obradjuju linearnu algebru, jednu od osnova modernog masinskog ucenja.

Tokom godina razvijeni su nekoliko kvantnih algoritama koji resavaju probleme linearne algebre. Zajedno ti algoritmi se nazivaju **osnovni kvantni**

**podprogrami linearne algebre** (eng. qBLAS), i oni se koriste u izradi algoritama za kvantno masinsko učenje.

Primeri algoritama koji su deo qBLAS-a su:

- HHL algoritam: koristi se za resavanje sistema linearnih jednačina, koristeći  $2^n$  dimenzionalni vektorski prostor za resavanje sistema sa  $n$  promenljivih. [1]
- Kvantna Furijeova transformacije [5]
- Kvantna procena faza za eigen vrednosti i eigen vektora/stanja. [5]

Ovi algoritmi se koriste kao osnova naprednijih algoritama i algoritama za Kvantno masinsko učenje. Samo treba pripaziti kod pominjanja ovih algoritama, jer neki od njih koriste neke koncepte koji su samo teorijske prirode ili su tesko kreirani u realnom svetu (npr. QRAM).

### 3 Kvantna teorija kompleksnosti

### 4 Kvantno masinsko učenje

Kvantno masinsko učenje je spoj kvantnih racunara i masinskog učenja. U programima Kvantnog masinskog učenja koriste se kvantni algoritmi (npr. qBLAS algoritmi) kao deo metoda optimizacija slične klasičnim metodama masinskog učenja.

Prema vrsti podataka koji se obradjuju oblast mozemo da dalimo na dve podoblasti

1. Obrada klasičnih podataka na kvantnim masinama (**Masinsko učenje dopunjeno kvantnim racunarima** eng. Quantum-enhanced machine learning)
2. Obrada kvantnim podataka na kvantnim masinama

Problem kod obrade klasičnih podataka na kvantnim masinama jeste ucitavanje podataka u sistem, kao i citanje rezultata. Ovo dovodi da algoritnim sa teorijskim eksponencijalnim ubrzanjem, u realnom svetu budu dosta sporiji i fizicki zahtevniji (velicina kvatnog kola zna da poraste i na skalu oko  $10^{25}$  za jednostavnu implementaciju HHL algoritma). [1]

## 4.1 Quantum support vector machine

Jedan od nejednostavnijih primera metoda Kvantnog masinskog učenja jeste **Quantum support vector machine** (QSVM). Klasican SVM je metoda koja pronalazi optimalnu podelu hiper-ravni izmedju dva razlicita skupa podataka, tako da sa velikom verovatnocom svi podaci iz jednog skupa podataka ce se naci na jednoj polovini hiper-ravni. [1]

### 4.1.1 Klasican algoritam

Ova metoda odredjuje klase koristeći linearnu funkciju  $w^T x + b$ . SVD predviđa prvu klasu ako je izlaz funkcije je pozitivan, a predviđa drugu klasu je izlaz negativan. Posto kod vecina slucajeva odvojenost izmedju dve klase podataka nije linearzibilno odvojivo, sa SVM metodom koristi se i **Kernel metoda**.

Pronalazenje optimalne hiper-ravni se sastoji od minimizacije  $|w|^2/2$  u nejednacini  $y_j(w * x_j + b) \geq 1$  za svako  $j$ . Ovo minimizacija se moze uraditi, ako uvedemo Karush-Kuhn-Tucker mnozioca  $\vec{\alpha} = (\alpha_1, \dots, \alpha_M)$  i maksimizujemo ih nad Lagranzovoj funkcijom:

$$L(\vec{\alpha}) = \sum_{j=1}^M y_j \alpha_j - \frac{1}{2} \sum_{j,k=1}^M \alpha_j \alpha_k x_j x_k$$

Sa sledecim ogranicenjima  $\sum_{j=1}^M \alpha_j = 0$  i  $\forall j \leq M \ y_j \alpha_j \geq 0$ . Tako da, parametre za hiper-ravan se izvide kao:  $w = \sum_{j=1}^M \alpha_j x_j$  i  $b = y_j - w x_j$  (za one  $j$  gde vazi da  $\alpha_j \neq 0$ ). Mali broj  $\alpha_j$  je razlicitno od nule, takve promenljive se odnose na vectore  $x_j$  koji leze na ravni, ti vektori se zovu **Support vektori** [7]

Kernel metoda transformise podatke u prostor gde su dve klase linearno odvojive. Metoda se oslanja na to da se linearna funkcija moze zapisati iskljucivo kao dot produkt izmedju primera.

$$w^T x + b = b + \sum_{i=1}^m \alpha_i x^T x_i$$

Gde je  $x_i$  trening primer a  $\alpha$  je vektor koeficijenata. Ovako zapisivanje funkcije nam dozvoljva da zamenim  $x$  sa izlazom funkcije  $\phi(x)$ , a dot produkt sa funkcijom  $k(x, x_i) = \phi(x) * \phi(x_i)$ . Funkcija  $k$  se zove **kernel**, dok funkcija



$\psi$  je funkcija koja preslikava podatke iz jednog prostora u drugi. Operator  $\langle * \rangle$  predstavlja unutrašnji produkt ekvivalentno  $\phi(x)^T \phi(x_i)$ . [2]

Kada zamenimo dot produkt sa kernelom, funkciju predikcije mozemo da zapisemo kao

$$f(x) = b + \sum_i \alpha_i k(x, x_i)$$

Jedan od velikih mana kernel metode jeste cena evaluacije izlaza kernel funkcije je linarna u odnosu na broj trening primera, jer  $i$ -ti bi oznacavao clana  $\alpha_i k(x, x_i)$  kernel funkcije. [2]

Slozenost SVM je  $O(\log(1/\epsilon)M^2(N + M))$ , gde je  $\epsilon$  preciznost resenja,  $N$  je broj dimenzija prostora nad kojim radimo ,a  $M$  je broj trening primera.

Takodje krajnje resenje se je binarni klasifikator za neki vektor  $x$ :

$$y(x) = \text{sign}(\sum_{j=1}^M \alpha_j k(x, x_j) + b)$$

#### 4.1.2 Kvantni algoritam

Pretpostavimo da imamo metodu za treniranje(eng. Oracle) koja vraca norme  $|x_j|$ , labele  $y_j$  i kvanten vektore  $|x_j\rangle = \frac{1}{|x_j|} \sum_{k=1}^N (x_j)_k |k\rangle$ .

Bitno nam je za algoritam da ova metoda vraca podatke pod donjom granicom, da bi se kompleksost jezgra algoritma mogla iskazati. Koristeci evaluaciju inner produkt priprema se kernel matrica, moze se dobiti SVD algoritam kompleksnoscu  $O(\log(1/\epsilon)M^3 + M^2 \log(N/\epsilon))$  Kernel matrica je od velike vaznosti za reformulaciju algoritma kao funkciju kvadratnog troska. Uvodimo simplifikaciju za nejednakosti, tako sto uvocimo promenljivu  $e_j$  i koristimo osoboinu labela da  $y_j^2 = 1$

$$y_j(w \cdot x_j + b) \geq 1 \rightarrow (w \cdot x_j + b) = y_j - y_j e_j$$

Porod ove jednacine imamo i implicitan uslov Lagranzove funkcije da sadrzi taksanu (eng. penalty) promenljivi  $\gamma/2 \sum_{j=1}^M e_j^2$  gde definisana  $\gamma$  za relativne tezinu greske treniranja. Ako uzmemo parcijalno derivat od Lagranzove funkcije i eliminisemo promenljivu  $u$  i  $e_j$  dovodi do aprokcismaciju funkcije kvadratnog troska problema:

$$F \begin{bmatrix} b \\ \vec{\alpha} \end{bmatrix} \equiv \begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & K + \gamma^{-1} \mathbb{1} \end{bmatrix} \begin{bmatrix} b \\ \vec{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{y} \end{bmatrix}$$

Ovde  $K_{ij} = x_i^T \cdot x_j$  je simetricna kernel matrica,  $y = (y_1, \dots, y_m)$  kao i  $\vec{1} = (1, \dots, 1)$ . Matrica  $F$  je dimenzija  $(M+1) \times (M+1)$ . Dodatna dimenzija (red i kolona) se sastoji od jedinica, zbog offset-a  $b$ . Promenljiva  $\alpha_j$  ima ulogu odredjivanje distance od optimalnog resenja. Tako da na resenje, odnonsno pronalazenje promenljivih za SVM je oblika:

$$\begin{bmatrix} b \\ \vec{\alpha} \end{bmatrix} = F^{-1} \begin{bmatrix} 0 \\ \vec{y} \end{bmatrix}$$

U klasicnom algoritmu kompleksnost SVM sa funkcijom kvadratnog troska je  $O(M^3)$

U kvantnom algoritmu, zadatak je generisanje stanja  $|b, \vec{\alpha}\rangle$  koja opisuju hiper-ravan i onda klasifikuju stanja  $|x\rangle$ . U algoritmu, resavamo normalizovanu jednacinu  $\hat{F}|b, \vec{\alpha}\rangle = |y\rangle$ , gde je  $\hat{F} = F/\text{tr}F$  sa ogranicenjem  $\|F\| \leq 1$ . Klasa ce biti odredjenja kao verovatnoca uspeha pri swap testu izmedju  $|b, \vec{\alpha}\rangle$  i  $|x\rangle$ . Za efikasnost merenje algoritma, posebno izracunavanja interzne matrice, matrica  $\hat{F}$  mora da se razdvoji na jednostavne elemente. Tako da matrica  $\hat{F}$  moze da se razdvoji na sledece elemente  $\hat{F} = (J + K + \gamma^{-1}\mathbb{1})/\text{tr}F$ . Gde je matrica

$$J = \begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & 0 \end{bmatrix}$$

Takodje, za estimaciju faze pravimo formulaciju Lijeovog produkta  $e^{-i\hat{F}\Delta t} = e^{-i\gamma^{-1}\mathbb{1}\Delta t/\text{tr}F} e^{-iJ\Delta t/\text{tr}F} e^{-iK\Delta t/\text{tr}F} + O(\Delta t)$

Za njega vazi da ima dve eigen vrednosti oblika  $\lambda_{\pm} = \pm\sqrt{M}$ , a, istovetno, eigen stanja su oblika  $|\lambda_{\pm}\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm \frac{1}{\sqrt{M}} \sum_{k=1}^M |k\rangle)$ . Za matricu  $\gamma^{-1}\mathbb{1}$  dve eigne vrednosti su  $v_1 = 0$  i  $v_2 = \gamma^{-1}M$ . Sada mozemo da aproksimiramo fazu za  $e^{-i\hat{F}\Delta t}$ .

Prvi korak, Stanje  $|y\rangle$  moze da se transformise u eigen state  $|u_j\rangle$  matrice  $\hat{F}$ , koja ima eigen vrednost  $\lambda_j$ . Ono je obilka  $|y\rangle = \sum_{j=1}^{M+1} \langle u_j|y\rangle |u_j\rangle$ . Ako inicijalizujemo aproksimaciju eigen vrednosti na  $|0\rangle$ , i primenimo estimaciju faze nad stanjem dobicemo stanje blize pravoj eigen vrednosti:

$$|y\rangle |0\rangle \rightarrow \sum_{j=1}^{M+1} \langle u_j|y\rangle |u_k\rangle |\lambda_j\rangle \rightarrow \sum_{j=1}^{M+1} \frac{\langle u_j|y\rangle}{\lambda_j} |u_j\rangle$$

Drugi korak je da invertujemo dobijeno stanje eigen vrednosti, pozivajuci rotaciju stanja. Na kraju dobijamo novo stanje sa trazanim parametrima

SVM ( $C = b^2 + \sum_{k=1}^M \alpha_k^2$ )

$$|b, \vec{\alpha}\rangle = \frac{1}{\sqrt{C}}(b|0\rangle + \sum_{k=1}^M \alpha_k |k\rangle)$$

**Klasifikacije** Sada imamo trenirani model kvantnog SVM-a i zelimo da klasifikujemo stanje  $|x\rangle$ . Od stanja  $|b, \vec{\alpha}\rangle$ , koriscenjem metode za treniranje, konstruisemo stanje:

$$|\tilde{u}\rangle = \frac{1}{\sqrt{N_u}}(b|0\rangle|0\rangle + \sum_{k=1}^M \alpha_k |x_k\rangle|k\rangle|x_k\rangle)$$

Gde nam je  $N_u = b^2 + \sum_{k=1}^M \alpha_k^2 |x_k|^2$ . Pored ovoga konstruisemo i ulazno stanje  $|\tilde{x}\rangle$ :

$$|\tilde{x}\rangle = \frac{1}{\sqrt{N_x}}(|0\rangle|0\rangle + \sum_{k=1}^M |x\rangle|k\rangle|x\rangle)$$

Gde nam je  $N_x = M|x|^2 + 1$ . Konstruisemo dva nova stanja  $|\psi\rangle$  i  $|\phi\rangle$ ;  $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|\tilde{u}\rangle + |1\rangle|\tilde{x}\rangle)$  i  $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ . Merenjem swap testa, verovatnoca dobivanja pozitivne vrednosti je  $P = |\langle\psi|\phi\rangle|^2 = \frac{1}{2}(1 - \langle\tilde{u}|\tilde{x}\rangle)$ . Ovde unutrasnji produkt, odnosno  $\langle\tilde{u}|\tilde{x}\rangle = \frac{1}{\sqrt{N_x N_u}}(b + \sum_{k=1}^M \alpha_k |x_k||x| \langle x_k|x\rangle)$ , koji se obicno izracunava u  $O(1)$  na kvatnom racunaru. Ako hocemo preciznost  $\epsilon$ , treba da iteriramo kroz algoritam merenja  $O(P(1-P)/\epsilon^2)$  puta.

## 4.2 Quantum principal component analysis

## References

- [1] Pancotti N. et al. Biamonte J. Wittek P. “Quantum machine learning”. In: *Nature* 549 (2017).
- [2] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. Adaptive Computation and Machine Learning series. MIT Press, 2016. ISBN: 9780262035613. URL: <https://books.google.rs/books?id=Np9SDQAAQBAJ>.
- [3] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. “Quantum algorithms for supervised and unsupervised machine learning”. In: *arXiv preprint arXiv:1307.0411* (2013).
- [4] Dan C. Marinescu. *Classical and quantum information*. Academic Press, 2012. Chap. 8. ISBN: 9780123838742; 0123838746.
- [5] Dan C. Marinescu. *Classical and quantum information*. Academic Press, 2012. Chap. 5. ISBN: 9780123838742; 0123838746.
- [6] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.
- [7] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. “Quantum support vector machine for big data classification”. In: *Physical review letters* 113.13 (2014), p. 130503.
- [8] Vlatko Vedral. *Introduction to Quantum Information Science*. Oxford University Press, 2006.