

11.10. 2023 8:15 – tímové stretnutie s vedúcou

Bod 1. – rozdeliť sa: Vlastne to je princíp toho delenia a princíp nejakého projektového riadenia, že keď si každý bude expert na niečo, tak to urobíte viac, čiže sa musíte kvázi rozdeliť hej? Čiže 1 časť vás sa bude venovať ďalej tomu datasetu a časť quassaru. Že budete zháňať dáta. Okej. No a hneď z dát sa dopĺňa aj tam z toho testu tak aby 100 mali vyrovnaný datasheet. **Cieľom je mať vyrovnaný dataset - doplniť datasety aby sa neprekrývali.** Prípadne pozrite nejaký ďalší sú už niečo použiť, či už pozitívne a negatívne, **proste čo najviac látok** a potom budeme pracovať s týmto setom. A 2. časť a teda niektorí ešte budú pokračovať v hľadaní toho aby našli ďalšie **datasety na iný typ genotoxicity iný typ testu genotoxicity.**

Áno dobre. No taký fajn začiatok. No a to by bolo tak to, že z ďalší sa rozdeľte, kto má na čo sa skôr expertízu a čo koho viacej baví. Tak uh ďalšie, kto sa bude venovať práve tomuto datasetu a bude sa snažiť nájsť nejaké ďalšie nové látky do toho, aby sme **v podstate si vytvorili na začiatku 2. datasety** ako podľa mňa teda neviem ako vy, ale teda najprv najhoršia práca je to vytvoriť, lebo toto hľadanie ihly v kope sena je to najhoršie. Bola som na konferencii o quasare a presne tam mala jedna žena prednášku o tom, **že vraj aj 50 látok je na kvasar dost'. To je blbosť** hej kravina hej? Ale uh, častokrát sú testy, kde tých med alebo nová metodika a tých dát je tam málo, čiže potom je to kvázi ako keby na nás ako sa s tým pohráme tak aby sme dostali čo najpresnejšiu predikciu. Len tak ďalej **treba pozerat' na to, že aké sú tam už po tej chemickej stránke štruktúry** a možno to fokusovať na nejaký typ štruktúry, lebo je sakra rozdiel, že keď si zoberiete toho modelu len jednu látku a ten vám nemôžete porovnávať zrazu inú štruktúru, vám to hodí kravinu. Preto **treba pozerat' na to, aké tam máte látky a že či môžem nazvať ten kvasar ako všeobecný**, nie že to bude super len na nejaké polycyklické látky, lebo sú tam len polycyklické látky.

Teda hovorili o testovaní 50 látok s relatívne presnými výsledkami. Bolo však spomenuté, že tieto látky mali veľmi podobnú štruktúru, čo znamená, že neexistovala veľká rôznorodosť. To by naznačovalo, že existujú ďalšie možnosti pre rozšírenie experimentu. Ďalej sme spomínali, že sa používajú odlišné metódy pre testovanie kultúr, najmä pri väčšom lineárnom datasete. To by malo jednoduchšie metódy ako pri kvasaroch umožniť rôzne postupy. Toto závisí od dostupných zdrojov, pretože je skvelé, ak môžeme využiť ľudské bunky, ale náklady môžu byť obrovské. Ak by sme chceli experimentovať in vivo, museli by sme použiť živú látku, čo by mohlo byť nákladné a vyžadovalo by sa množstvo testov. V tomto prípade by sa pohybovali na sumách medzi 80 000 a 100 000 dolárov, ktoré by predstavovali veľký experimentálny rozpočet

Proteíny - tieto molekuly majú významný vplyv na naše myšlienky, pocity a náladu, a to či sme apatickí alebo pozitívne naladení. Regulujú všetko, od enzýmov, ktoré premenia jednu vec na druhú, až po proteíny, ktoré nesú informácie a signály. No, existujú aj proteíny, ktoré fungujú podobne, nesú hormóny, ako napríklad inzulín, ktorý reguluje hladinu cukru v krvi. No a potom tu sú ďalšie proteíny, ktoré sú dôležité pre rôzne stavy, ako napríklad depresia. Niektoré proteíny pomáhajú zabrániť starnutiu a udržiavať nás mladšími, napríklad sirtuíny. Mutácie v DNA sú bežné, dokonca sa nám každý deň vyvinie tisíce až státisíce. Problém vzniká neskôr, okolo 30. roku života, keď sa naše opravné mechanizmy začínajú prejavovať. Niektoré proteíny vedia opravovať tieto mutácie a brániť tak rôznym chorobám. Genetika a životný štýl hrajú dôležitú

úlohu. Existujú aj látky, ktoré môžu stimulovať tieto proteíny, ako napríklad metformín, ktorý sa používa na liečbu diabetu a súčasne zastavuje starnutie. Takže v konečnom dôsledku je všetko spojené s proteínmi.

Genotoxicita označuje, čo sa stane, keď sme vystavení určitej látke, buď exogénnej, čo znamená, že môže byť čímkoľvek, či už ide o chemickú látku nachádzajúcu sa prirodzene v citróne alebo chemickú zložku vytvorenú na účely farbenia kozmetiky alebo iným spôsobom. Tieto látky sa nejako spájajú s našou DNA, čo má potenciálne negatívne dôsledky. Naším cieľom je skúmať tieto látky, z ktorých existuje mnoho, a zistiť, či ich dlhodobé vystavenie môže mať nejaký škodlivý vplyv. Toto je obzvlášť dôležité v súvislosti s genotoxickými účinkami na tehotenstvo.

Keď sú ženy tehotné alebo dojčia, aj malé množstvo genotoxických látok môže spôsobiť vážne poškodenie plodu. Preto je nesmierne dôležité monitorovať genotoxicitu. Toto je jedna z hlavných oblastí, na ktoré sa regulačné orgány pýtajú povolenia na použitie rôznych látok, nielen bežne dostupných, ale aj liekov. Často sa stáva, že aj lieky, ktoré sú bežne používané, podliehajú testom genotoxicity, pretože aj pri ich krátkodobom užívaní môžu spôsobiť poškodenie. Regulačné orgány vyžadujú tieto testy, a ako alternatívu k tradičným testom sa začínajú prijímať modely, ktoré môžu určiť genotoxicitu látok bez potreby testov na živých organizmoch. Týmto spôsobom sa snažíme skrátiť a zjednodušiť proces schvaľovania nových látok, ktoré majú prísť na trh a zabezpečiť, že nebudú škodlivé pre ľudí ani životné prostredie. Genotoxicita hrá kľúčovú úlohu v regulácii bezpečnosti látok.

Takže poďme sa ešte raz vrátiť k tým datasetom. Máte tento dataset, s ktorým budete pracovať.

Vašou úlohou je získať z týchto dát látky, ktoré sa navzájom čo najviac podobajú.

Neznamená to nutne rovnaký počet látok v každej skupine, ale **snažte sa dosiahnuť čo najvyváženejší pomer medzi pozitívnymi a negatívnymi prípadmi**. Napríklad, aj keď máte 10 pozitívnych a 40 negatívnych vzoriek, to je v poriadku.

Druhá časť vašej práce bude spočívať v hľadaní relevantných in vivo testov. Bude to dôležité, pretože výsledky týchto testov budú mať výrazný vplyv na vašu modelovaciu prácu. Je zaujímavé, že **tieto testy môžu viesť k vytvoreniu úplne odlišného modelu**, a to bude veľmi zaujímavý aspekt vašej práce.

Budúci týždeň je konferencia. Už sme prešli, čo by sme mali urobiť v prvom kroku, a teraz je potrebné zvážiť nasledujúce kroky. V prípade, že by som mala problémy s pripojením na internet, pretože sa konferencia koná v Tunisku, a neviem, ako to bude so signálom, môžem vám zavolať a dohodnúť sa na ďalších krokoch. Môžem vám oznámiť postup, ak by sa tak stalo.

Jedným z ďalších krokov je **nájsť čo najviac molekulárnych deskriptorov, ktoré sú dostupné v súvislosti s RTG c-ky. Na túto úlohu môžeme použiť nástroje, ako sú RDKit alebo CDKit.** - Tooly na rátanie deskriptorov rdkit cdkit.

Čiže **deskriptory** môžu zahŕňať molekulovú hmotnosť, počet uhlíkov v molekule, počet dusíkatých atómov, počet sírových atómov a iné parametre. Môžu to byť charakteristiky spojené s chemickými vlastnosťami molekúl, ako napríklad rozpustnosť alebo biodostupnosť. Taktiež

môžu obsahovať parametre týkajúce sa štruktúry molekuly, ako je počet cyklov v molekule a ďalšie. Nasledujúcim krokom je vypočítať tieto deskriptory na vašom datasete a na základe týchto charakteristík začneme analyzovať ich vplyv. Potom sa už rieši, že ktorý je najlepší.

Dobre uh tie **sprinty** – v skratke, budú **dvojťždňové**.

Potom si musíte **vytvoriť webovú** stránku pre tímak. Musíte sa ešte dohodnúť, kto bude mať na starosti jej tvorbu a tom, ako to bude vyzerat'. Momentálne nie je nič definitívne, takže nie je jasné, aký bude konečný výsledok.

Teda **vlastne v tomto zimnom semestri, skôr by sme sa mali hrať s tými dátami** a potom v letnom semestri na tu webovú aplikáciu. Hej, skúsme spraviť aby nám to fičalo a potom už sa budeme hrať, budeme zdokonaľovať trochu model.

11.10. 2023 20:30 – tímové stretnutie

Preberané témy:

- rozdelenie úloh - bude prebiehať na báze konsenzu, pravdepodobne
- ako budeme hľadať datasety - postup: nájsť, dať na schválenie