



MIM SOLUTIONS

Rekomendacja terapii dla INVICTA

Zbiory danych

- ★ Zbiór **kilkuset procesów stymulacji hormonalnej** z informacjami z wynikami badań genetycznych
- ★ Zbiór kilkunastu tysięcy procesów stymulacji hormonalnej bez danych genetycznych

Kolumny

- ★ Zbiór danych z kilkaset kolumnami w kilku plikach, znaczna część nieistotnych z punktu widzenia **weryfikowanego rezultatu, a więc efektów stymulacji**
- ★ Blżej przyjrzelśmy się około **100** kolumnom
- ★ W zbiorze polimorfizmów mamy **ponad trzydzieści tysięcy** wpisów dotyczących **kilkuset** różnych wariantów
- ★ Z tego wyselekcjonowaliśmy listę **kilkudziesięciu** wariantów podejrzanych o jakikolwiek wpływ, którą następnie zawęziliśmy do kilku

Filtrowanie danych

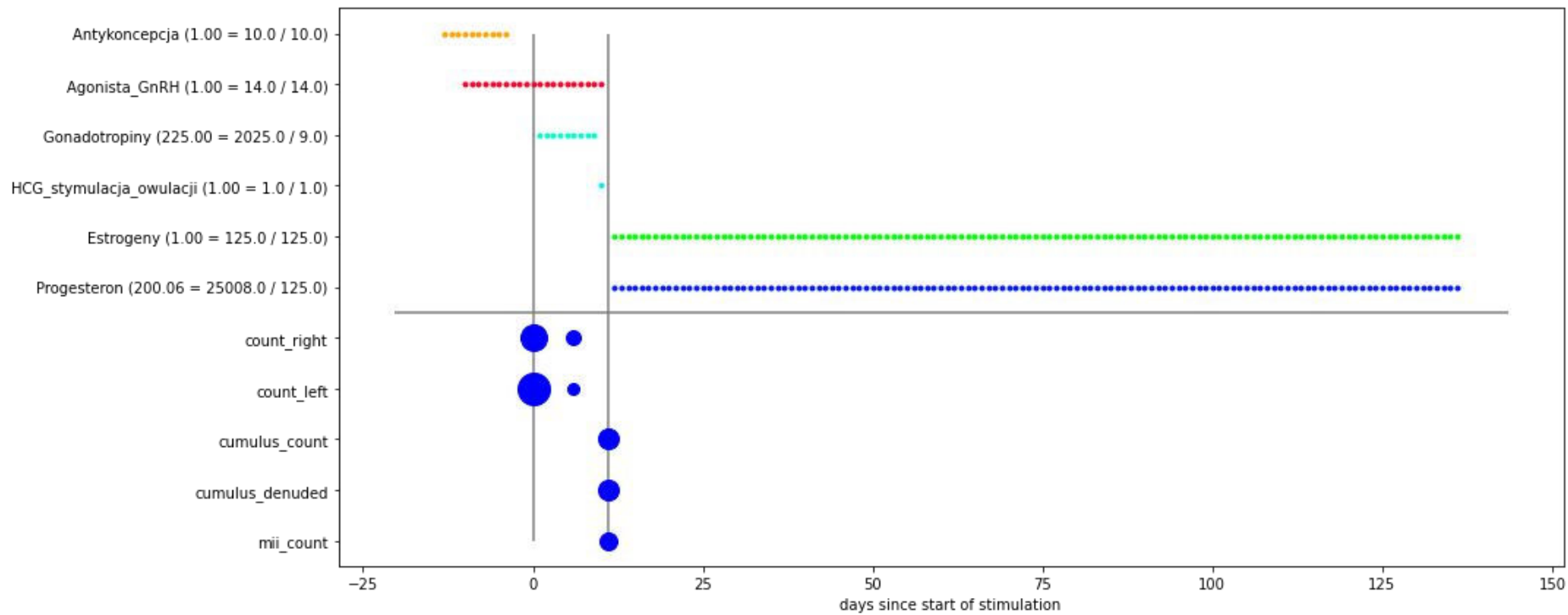
Z analizy wyeliminowano:

- ★ Dawczyni komórek, biorczynie komórek, dodatkowe stymulacje,
- ★ Procesy pacjentek o AMH > 15.0.

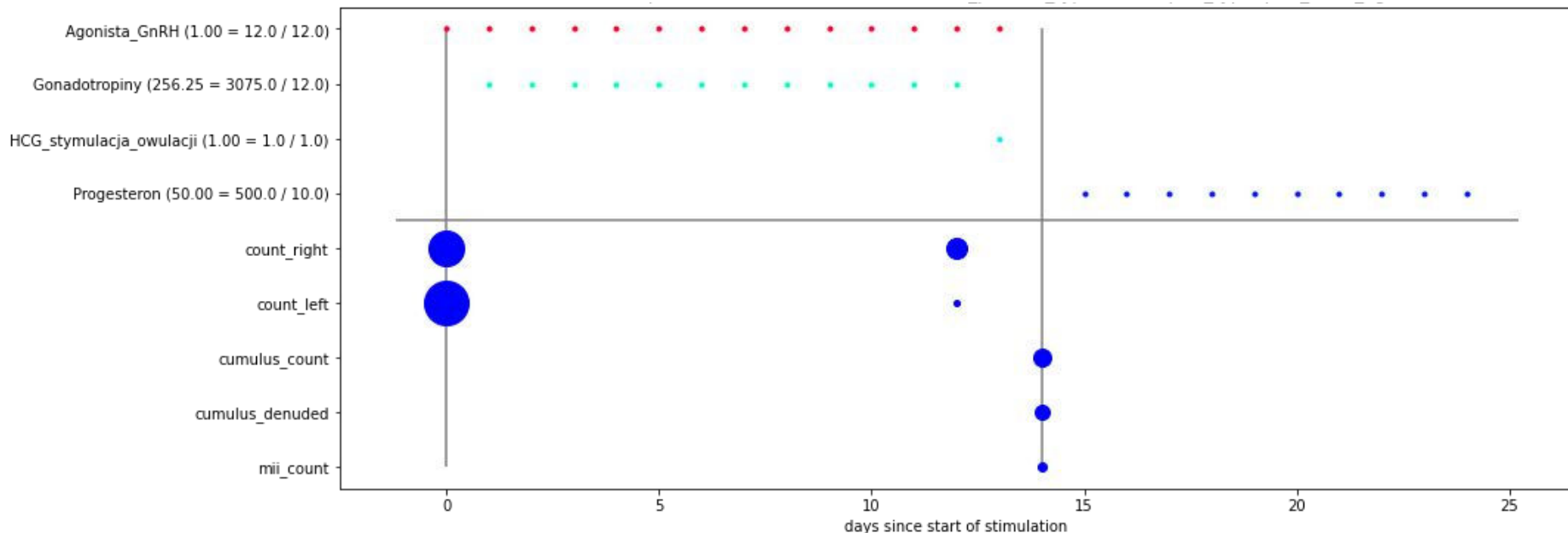
Podstawowe dane (wyniki badań, itd.)

- ★ Wiek pacjentki
- ★ Wyniki badań hormonalnych (m. in. AMH)
- ★ Przyczyny niepłodności
- ★ Polimorfizmy genetyczne
- ★ Wyniki badań liczby pęcherzyków z pierwszego dnia stymulacji
- ★ Wyniki z dnia pickupu: liczbę uzyskanych cumulusów i dojrzałych komórek jajowych

Przykład procesu długiego



Przykład procesu krótkiego z agonistą



Sytuacje w których można użyć modelu

- ★ Do rekomendowania **typu protokołu** [dzień decyzji o rozpoczęciu terapii]
- ★ Do rekomendowania dawek gonadotropin [start symulacji]

Co przewiduje i optymalizuje model?

- ★ Przewidywany efekt stymulacji hormonalnej: liczbę komórek jajowych, którą można uzyskać
- ★ Prawdopodobieństwo hiperstymulacji

Modele bazowe

- Model **BASE_0** oparty na średniej
- Model **BASE_1** oparty tylko na **AMH**
- Model **BASE_2** oparty na kolumnach:
 - AMH
 - wiek pacjentki
 - Liczba pęcherzyków antralnych na pierwszy dzień stymulacji
 - liczba cumulusów w poprzednim procesie - jeśli taki był
 - liczba dojrzałych komórek jajowych w poprzednim procesie - jeśli taki był
- Model **BASE_3** oparty dodatkowo na 7 wyselekcjonowanych polimorfizmach

Planowane funkcjonalności systemu

- ★ Dla każdej decyzji (np. wybór protokołu) określamy oczekiwaną liczbę dojrzałych komórek jajowych oraz ryzyko hiperstymulacji, sugerujemy optymalny wybór.
- ★ System działa w dwóch wersjach z pakietem genetycznym oraz bez pakietu.
- ★ [Opcjonalnie] Określenie najważniejszych czynników, które mają wpływ na rokowania pacjentki.

Podsumowanie prac

- ★ Przygotowanie danych (wczytanie, zmergeowanie, odfiltrowanie błędnych)
- ★ Framework do trenowania modeli uwzględniający:
 - Dwa zbiory danych
 - Różne cele: cumulus_denuded, mii, certainty_5%, certainty_95%, hiper_20, hiper_25
 - Różne miary wyników: RMSE, MSE, MAPE
 - Cross-walidację wyników
- ★ Framework do uruchamiania eksperymentów i porównywania wyników
- ★ Wyselekcjonowanie kolumn istotnych dla modelu
- ★ Interpretacja modelu: SHAP
- ★ Selekcji genów algorytmem Boruta
- ★ Implementacja analizy counterfactualnej
- ★ Analiza wariantów
- ★ Algorytmy typu KNN.

Kolumny używane w eksperymentach

- ★ Przyczyny niepłodności
- ★ Hormony
- ★ Hormony - unormowane (w normie / poniżej / powyżej)
- ★ Hormonów - z uwzględnieniem jedynie odpowiednio świeżych badań
- ★ Polimorfizmy genetyczne
- ★ Wyniki badań USG z ds1
- ★ Typ protokołu
- ★ Dawki gonadotropin
- ★ Dawki i czas przyjmowania pozostałych leków
- ★ Wyniki poprzednich procesów danej pacjentki (jeżeli były)
- ★ Oraz różne kombinacje pomiędzy powyższymi