

Taksówka już po Ciebie jedzie!

...czyli machine learning w biznesie.

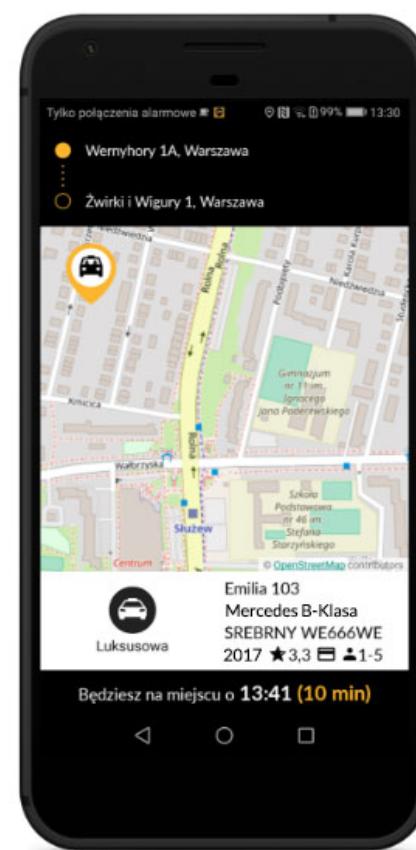
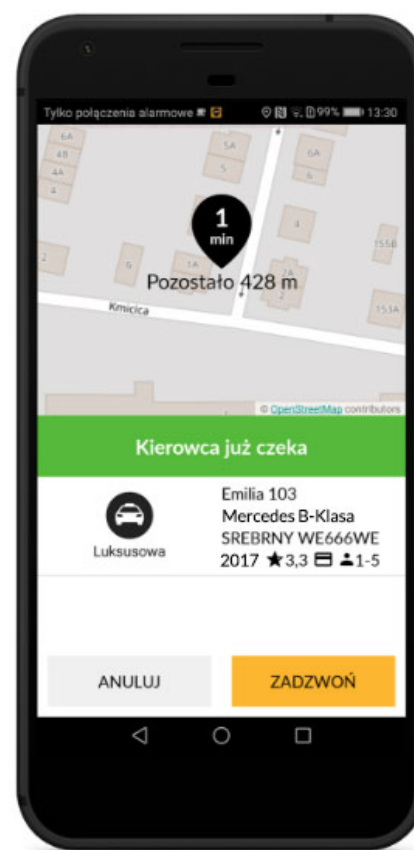
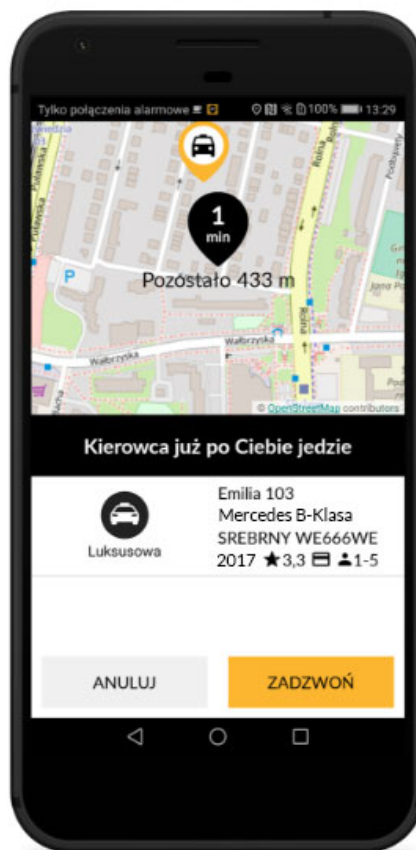
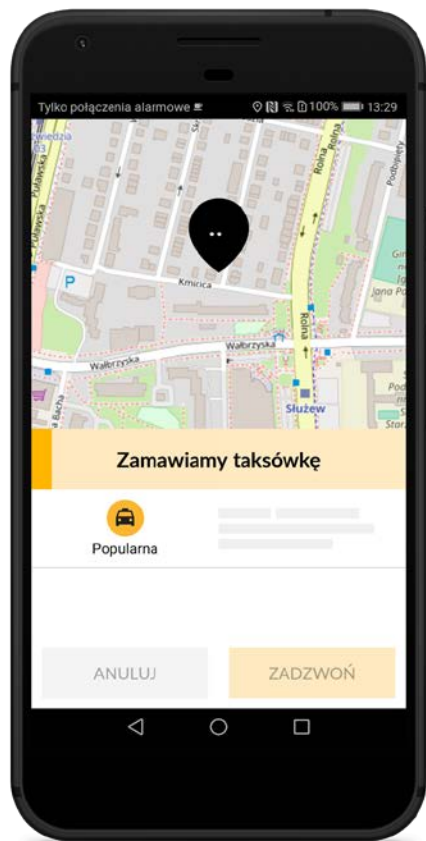
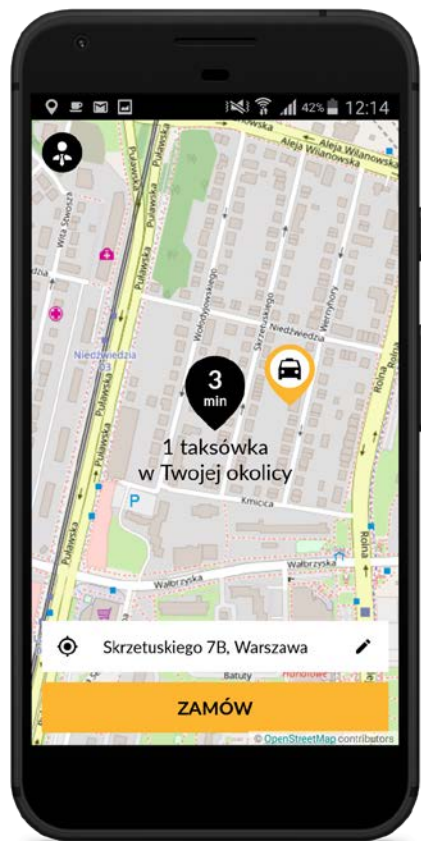
Historia wdrożenia dynamicznego algorytmu w firmie iTaxi.

Tomasz Brzeziński
Chief Data Scientist



Aplikacja Taxi

Jak działa iTaxi?

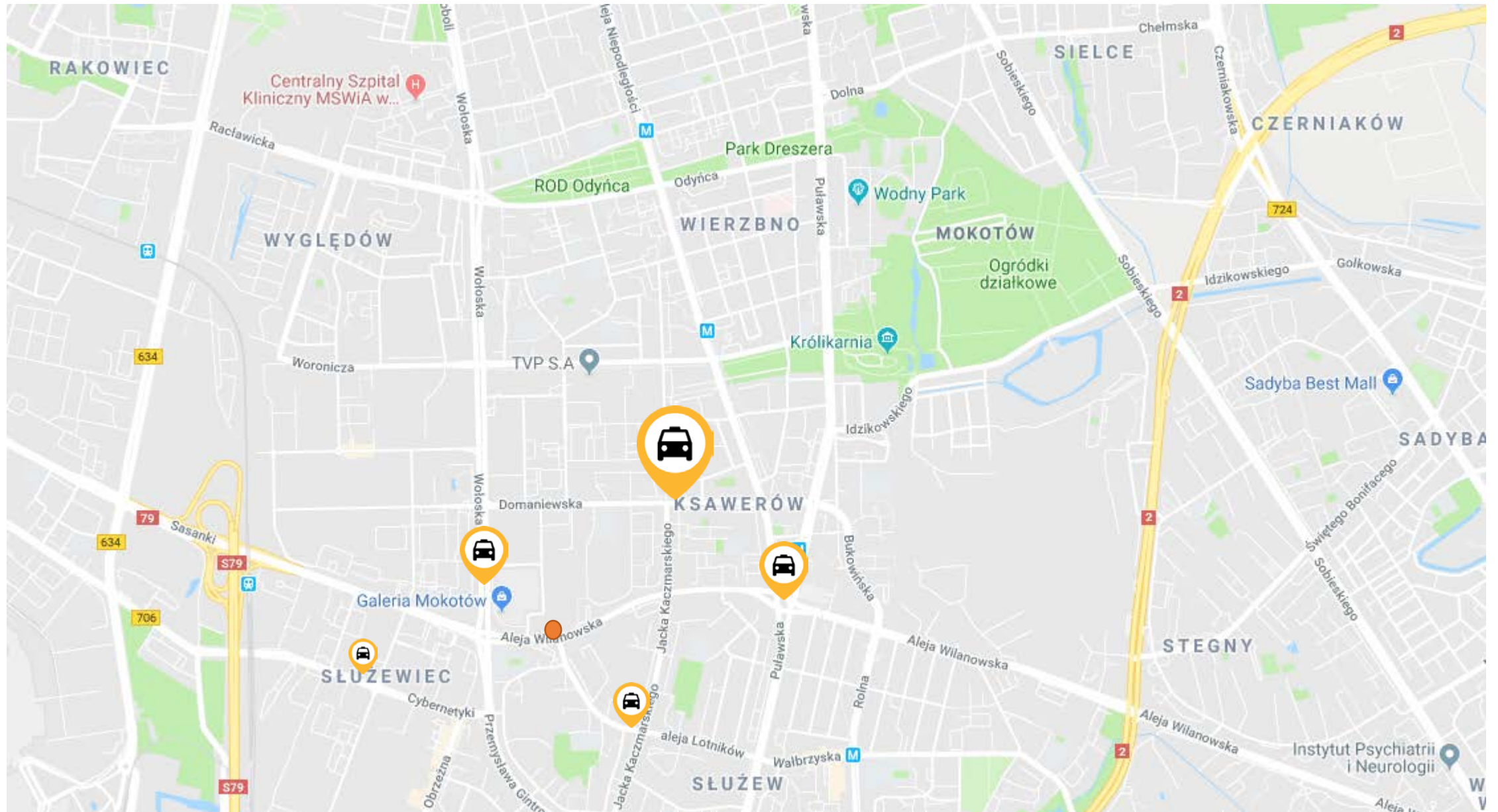


Problem: optymalizacja algorytmu wydawania
zleceń



Aplikacja Taxi

Którą taksówkę wysłać?



Dostarczyć taksówkę:

- jak najszybciej
- jak najlepszą
- z żądaną stawką za kilometr
- spełniającą dodatkowe wymagania (np. liczba osób)
- tak, aby móc skutecznie zrealizować kolejne zlecenia

Dostarczyć taksówkę:

- jak najszybciej: **mediana czasu dojazdu**
- jak najlepszą: **średni scoring kierowcy**
- tak, aby móc skutecznie zrealizować kolejne zlecenia:
reliability ratio (ilość kursów/popyt)

Dlaczego to jest trudne?

Problem

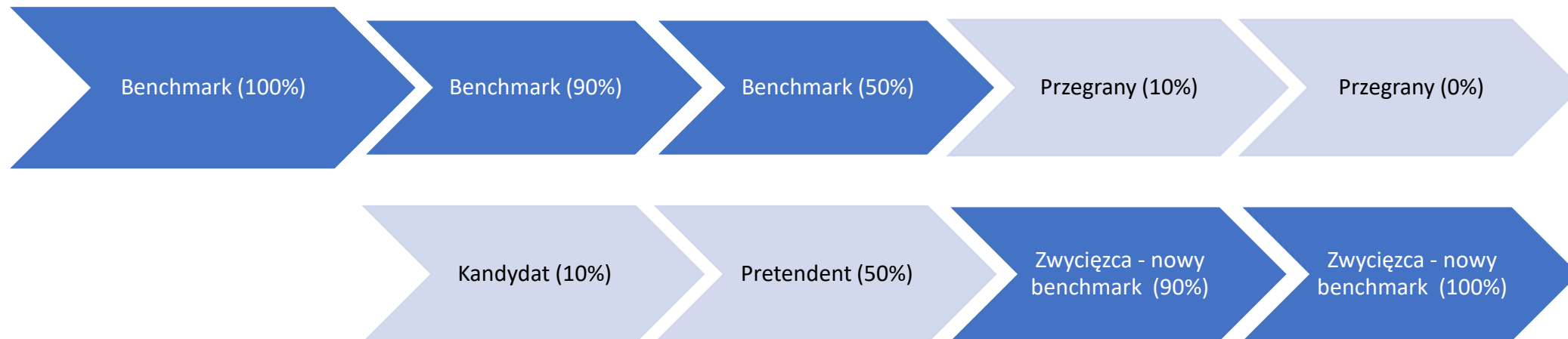
- wielokryterialny
- nieliniowy i niesprowadzalny do liniowego
- dyskretny
- niestabilny w czasie
- w warunkach niepewności

System cyklicznej poprawy algorytmu

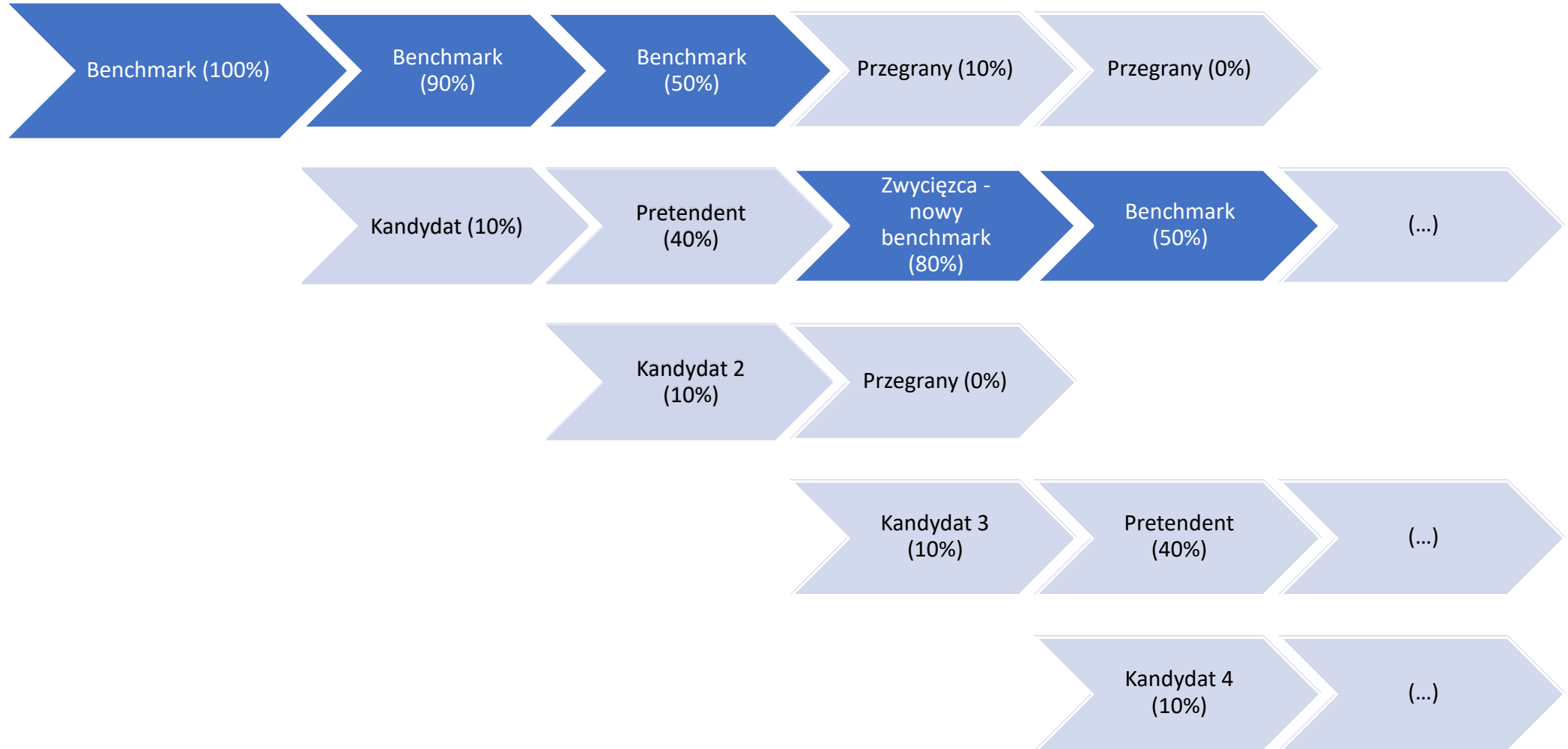


Aplikacja Taxi

Testy A/B



Cykliczne testy A/B



Ale...

„kto psuje jakąś rzecz, żeby lepiej poznać jej istotę, ten zbacza ze ścieżek mądrości”



Aplikacja Taxi

Wielki Zderzacz Taksówek

...czyli symulator miasta



Aplikacja Taxi

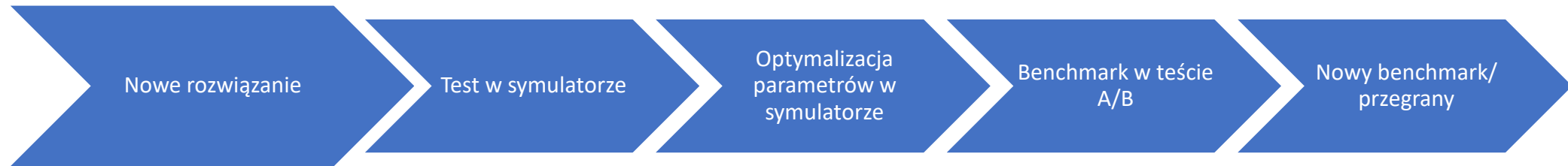
Odwzorowanie zachowania miasta:

- Taksówkarze zaczynają i kończą pracę (w różnych godzinach)
- Pasażerowie zamawiają taksówki
- Algorytm wybiera najlepszą taksówkę
- Taksówkarz przyjmuje zlecenie lub odrzuca
- Pasażer może zrezygnować
- Taksówkarz jedzie po pasażera
- Pasażer wsiada
- Kurs jest realizowany

Symulator - technikalia

- Kod w Pythonie (IDE: pycharm)
- Około 3000 linii kodu (3 tygodnie mojego życia)
- Graficzne przedstawienie tylko dla debugowania
- Uruchamiany lokalnie
- Efektywność: 1 dzień symulacji dla Warszawy – ok. 3,5 sekundy (dzięki wektoryzacji przekształceń)
- Optymalizacja polega na testowaniu różnych wersji algorytmu przydziału taksówki na danych historycznych

Pretesting w symulatorze

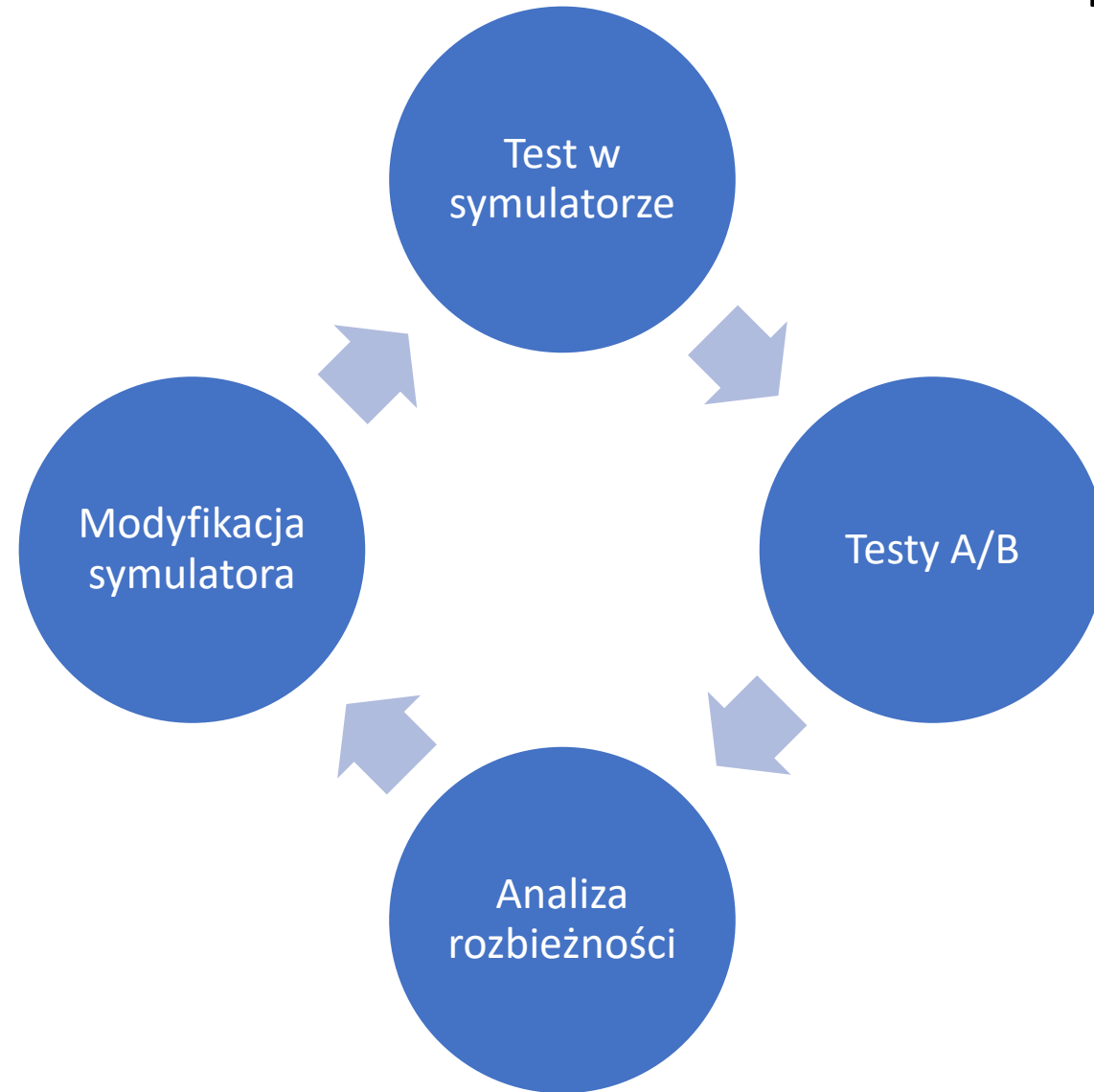


A jeśli wyniki w rzeczywistości odbiegają od symulacji?



Aplikacja Taxi

Kompletny system



Symulator - wersja 1



Aplikacja Taxi

Odwzorowanie zachowania miasta:

- Taksówkarze zaczynają i kończą pracę (w różnych godzinach)
- Pasażerowie zamawiają taksówki
- **Algorytm wybiera najlepszą taksówkę**
- Taksówkarz przyjmuje zlecenie lub odrzuca
- Pasażer może zrezygnować
- Taksówkarz jedzie po pasażera
- Pasażer wsiada
- Kurs jest realizowany

Odwzorowanie zachowania miasta:

- Taksówkarze zaczynają i kończą pracę (w różnych godzinach)
- Pasażerowie zamawiają taksówki
- Statyczny (zależny od parametrów znanych a priori) algorytm wybiera najlepszą taksówkę
- Taksówkarz przyjmuje zlecenie lub odrzuca
- Pasażer może zrezygnować
- Taksówkarz jedzie po pasażera
- Pasażer wsiada
- Kurs jest realizowany

Wyniki



Aplikacja Taxi

- Zmiana średniego scoringu: +0,4%
- Zmiana mediany czasu dojazdu: -8,5%
- Zmiana RR: +7,5%

Dlaczego tak mało?



Aplikacja Taxi

Dlaczego tak mało?

- Rzeczywistość nieustannie się zmienia
- Optymalne algorytmy są specyficzne dla miast i czasu
- Testy A/B muszą być ostrożne -> określenie istotności statystycznej zajmuje czas

Symulator - wersja 2



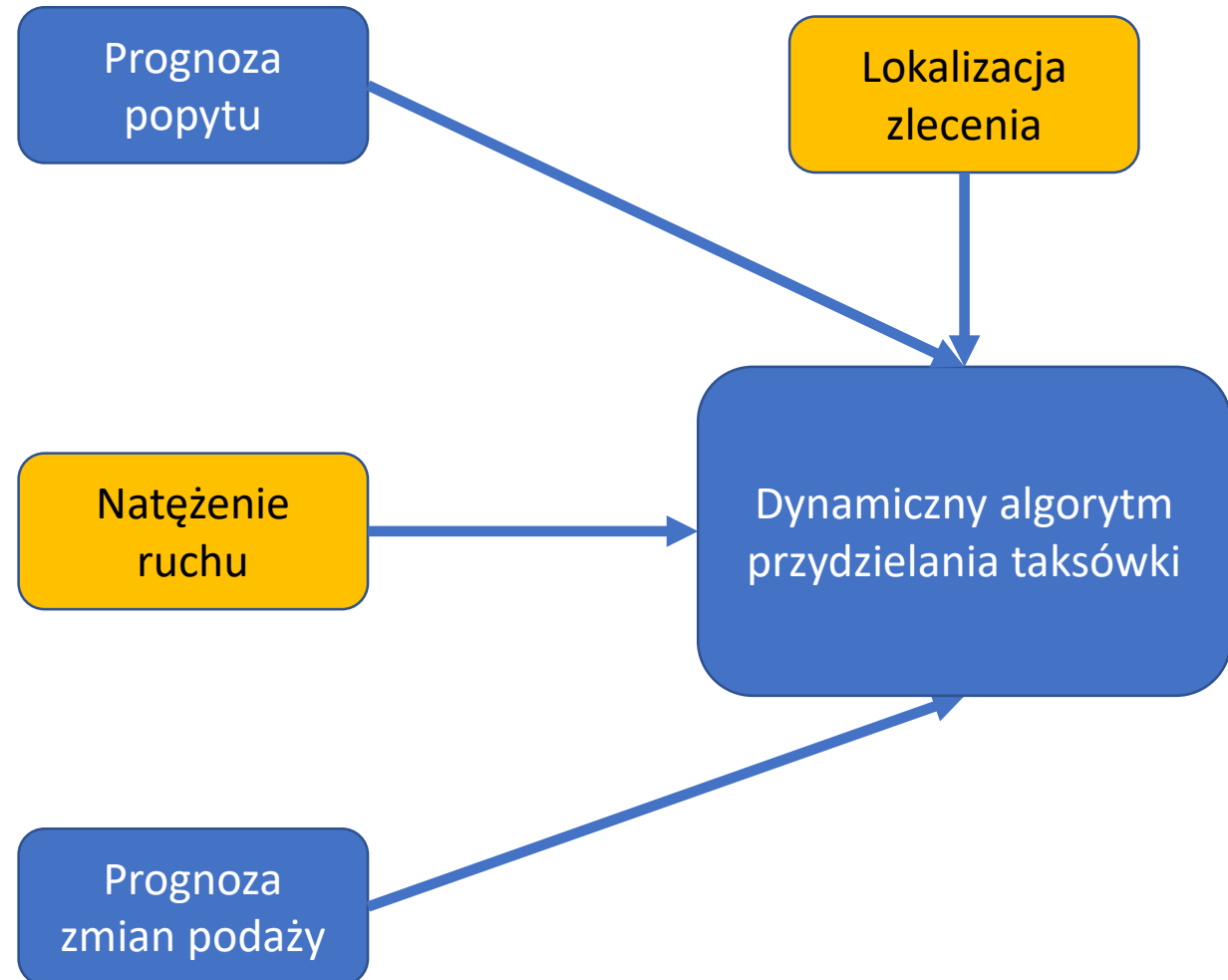
Aplikacja Taxi

Odwzorowanie zachowania miasta:

- Taksówkarze zaczynają i kończą pracę (w różnych godzinach)
- Pasażerowie zamawiają taksówki
- **Dynamiczny (uwzględniający stan miasta) algorytm wybiera najlepszą taksówkę**
- Taksówkarz przyjmuje zlecenie lub odrzuca
- Pasażer może zrezygnować
- Taksówkarz jedzie po pasażera
- Pasażer wsiada
- Kurs jest realizowany

Wersja 2 - dynamiczna automatyczna optymalizacja

- Prognoza popytu i podaży jest tworzona dynamicznie dla najbliższego otoczenia punktu zlecenia
- Anomalie (korki, zachwianie równowagi między popytem i podażą) są w krótkim czasie automatycznie uwzględniane przez algorytm



Wyniki



Aplikacja Taxi

- Zmiana średniego scoringu: $+0,4\% \Rightarrow +0,3\%$
- Zmiana mediany czasu dojazdu: $-8,5\% \Rightarrow -12,4\%$
- Zmiana RR: $+7,5\% \Rightarrow +8,2\%$

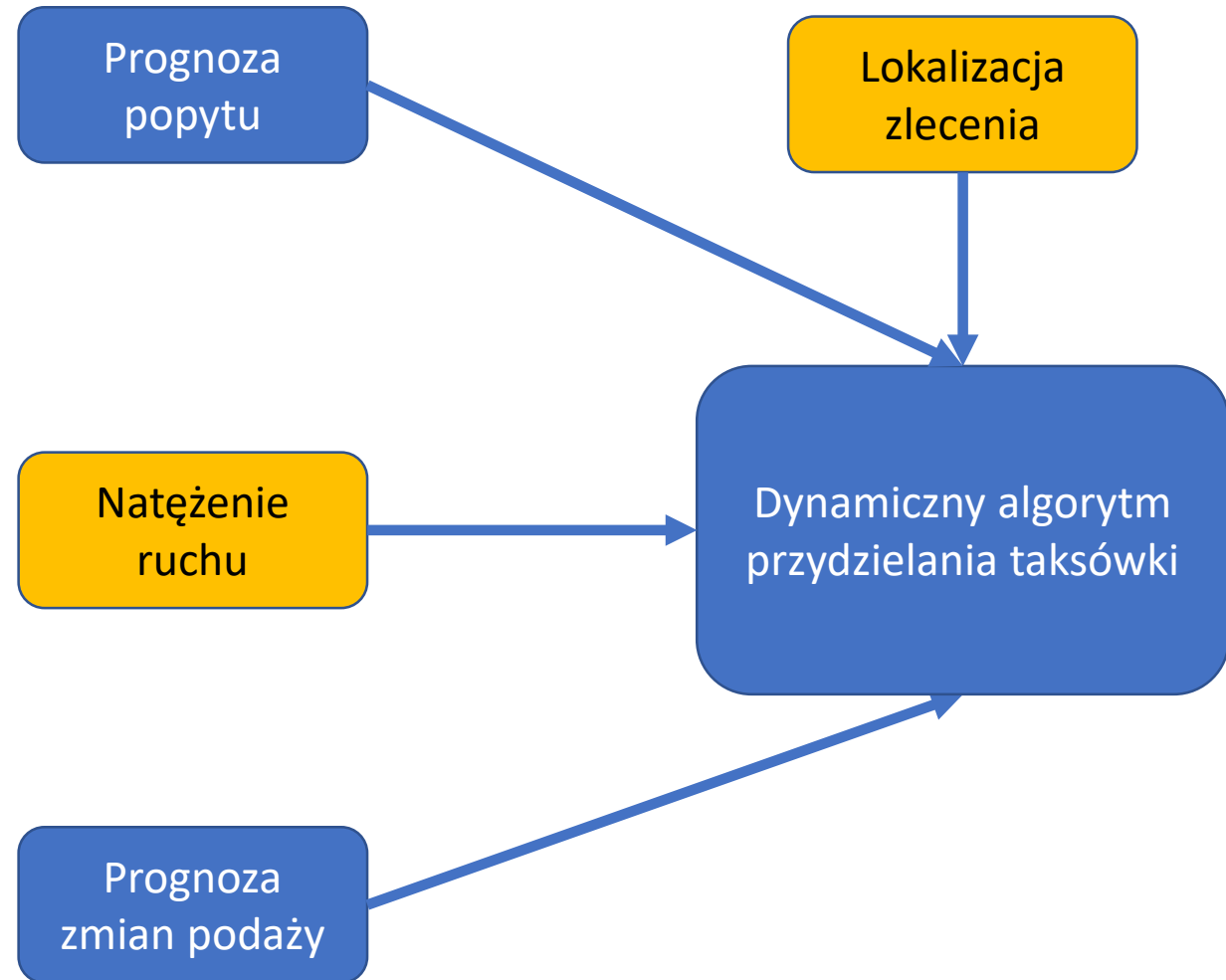
Symulator - wersja 3
(w budowie)



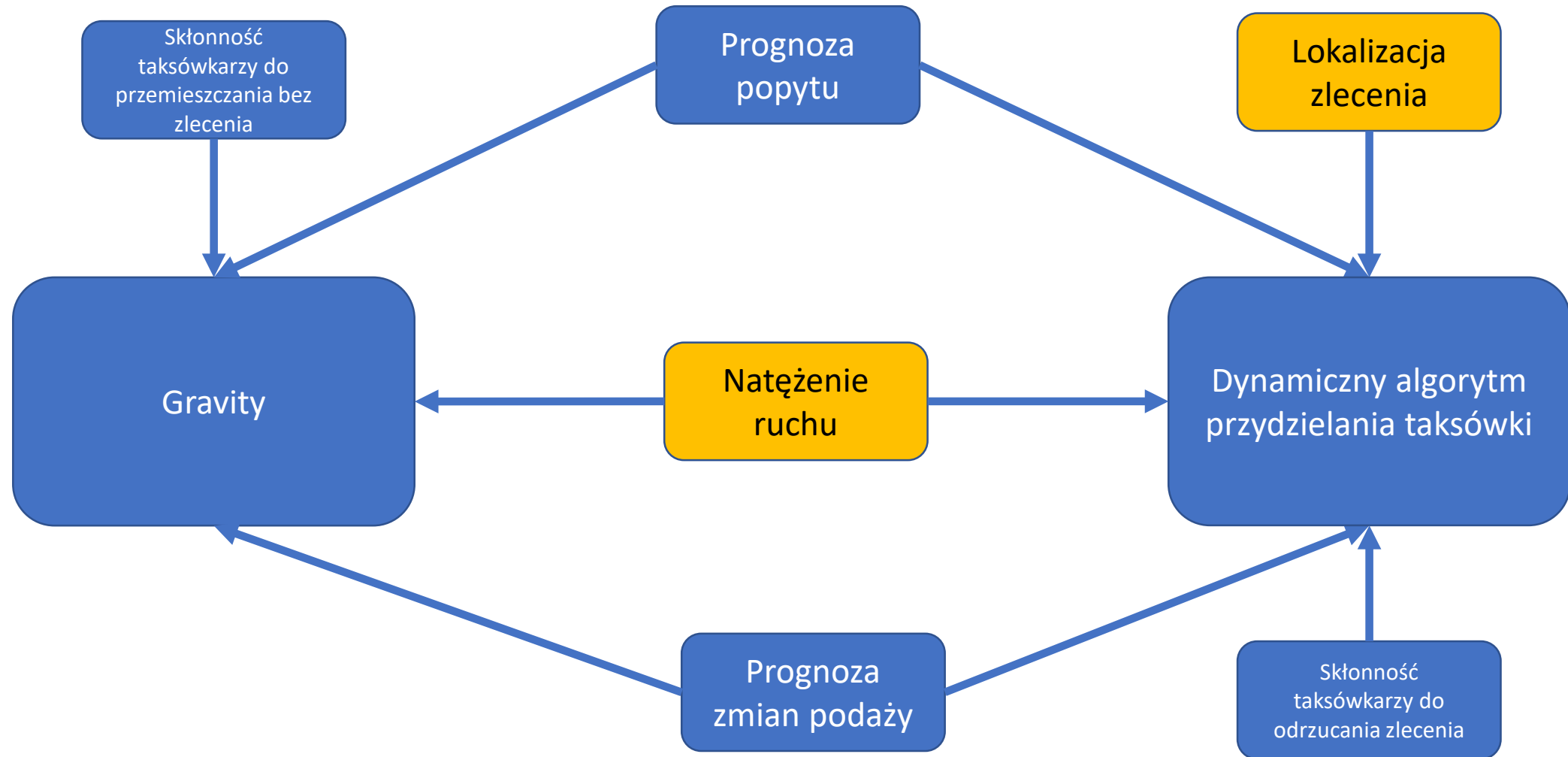
Aplikacja Taxi

- Przewidywanie, gdzie pojawią się zlecenia
- Sugerowanie taksówkarzom, gdzie powinni się przemieścić, aby zwiększyć swoje szanse na zdobycie kursu

Wersja 2 - dynamiczna automatyczna optymalizacja



Wersja 3 – dynamiczna optymalizacja + gravity



Efekty uboczne



Aplikacja Taxi

Dodatkowe korzyści

- Ułatwione wykrywanie fraudów (i innych anomalii)
- Środowisko do testowania nowych rozwiązań
- Lepsze zrozumienie (dość skomplikowanego) biznesu
- System wczesnego ostrzegania przed zmianą otoczenia biznesowego

Na wynos

- Jeśli możesz, przed testowaniem na żywo sprawdź scenariusz w warunkach laboratoryjnych
- Jeśli problem jest zbyt skomplikowany do rozwiązania analitycznego, przeprowadź symulację
- Rozbieżności wyników symulacji z wynikami testów A/B są narzędziem do doskonalenia symulatora
- W zmieniającej się rzeczywistości optymalizacja nigdy się nie kończy