

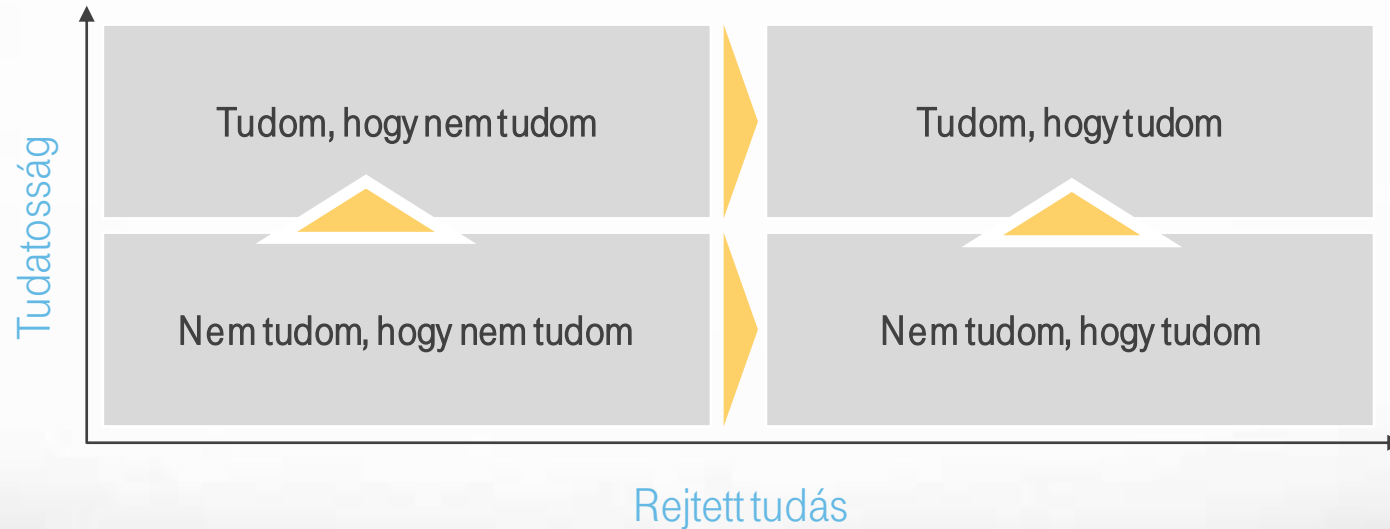
KÜLÖNBSÉG A SZABÁLY ALAPÚ ÉS AZ ALGORITMIKUS ATTRIBÚCIÓS MODELLEK KÖZÖTT

NAGY BARNABÁS

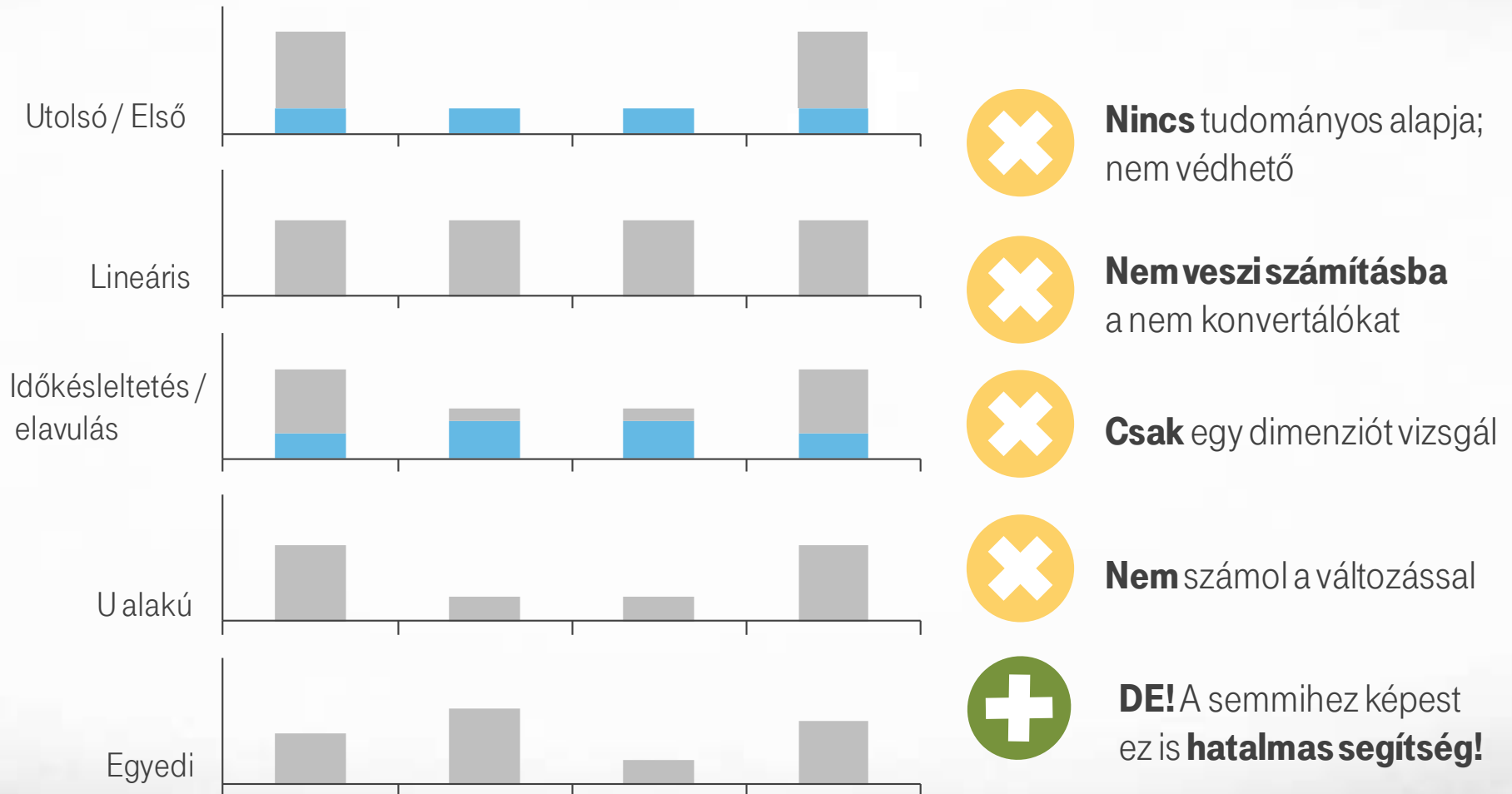
MAGYAR TELEKOM
E-COMMERCE ÉS E-CRM OSZTÁLY

ÜZLETI NÉZET

- **Szabály alapú modellek:** hipotézist vizsgál és tudást pontosít - kicsiben is elkezdhető
- **Algoritmus alapú modellek:** Rejtett tudást is feltár, amire nem volt hipotézis , emberileg nem követhető nagyságú dimenzió és sebességű számításokon alapul - nagy lépés



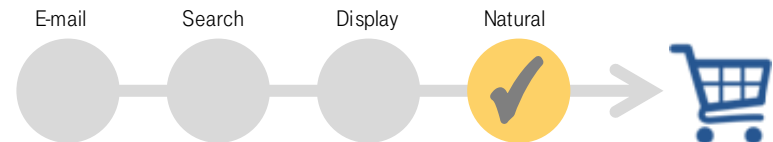
A SZABÁLY ALAPÚ MODELLEK ÉS HIÁNYOSSÁGAIK



A SZABÁLY ALAPÚ MODELLEK ÉS HIÁNYOSSÁGAIK MÁS SZEMPONTOKBÓL

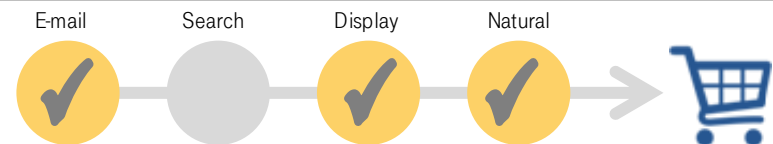
Utolsó kattintás

- **Hiányosság:** Csak az utolsó kattintást értékeli



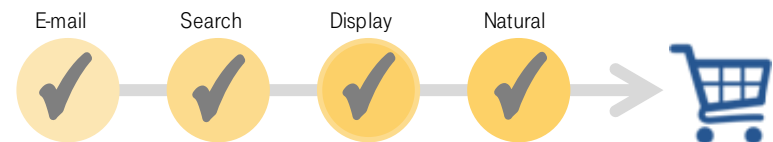
Egyenletes attribúció

- **Hiányosság:** Még túlbecsüli a frekvenciát



Időkésleltetés /elavulás

- **Hiányosság:** Feltételezi, hogy minden csatorna ugyanolyan értékkel bír

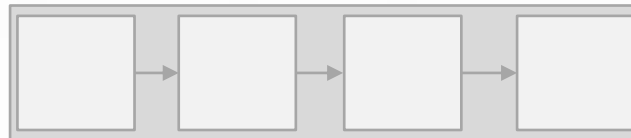


A szabály alapú modellek nem számolnak:

Szezonalitással, hirdetői trenddel, szaturációval, makroökonómiai hatásokkal, eszközök közötti aktivitással vagy a cookie elvesztéssel, törléssel

ALGORITMUS ALAPÚ MODELLEK ÉS ELŐNYEIK

1. ÚTVONAL (PATH)



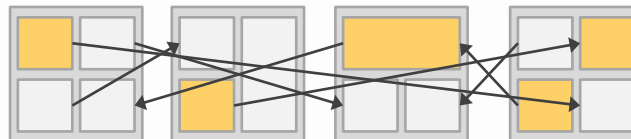
2. MULTI DIMENZIONÁLIS



Attribútum súlya



Interaktív
többváltozós funkció



Attribútum fontossága



Normalizálás & attribúció



AZ ALGORITMUS ALAPÚ MODELLEK

1. Path és top-down megközelítés

Távolsági (Altitude) Modellek

Top-Down modell



Útvonal szintű modell

Cél

- Elkülöníti az inkrementális és alap konverziókat
- "Lift" index számítás csatornánként (hatékonyság index)
- Becsült elavulási rátát csatornánként
- Számol a cookie veszteséggel

Inputok

- Napi összesített csatorna és konverziós adat
- Gazdasági adat (pl. S&P 500)
- Hirdetői trend
- Vertikális szezonális
- Árazás és kedvezmények

Cél

- Értelmezi az útvonal sorrendjét és dinamikáját
- Újszerűség (elavulás)
- Frekvencia
- Markov-lánc dinamika

Inputok

Az útvonal szintű adat tartalmazza:

- Csatorna érintkezési pont az útvonalon
- Esemény típus
- Útvonal struktúra
- Időbélyeg
- Forrás, kampány
- Frekvencia

A két modellt kombinálva

- A távolsági modell kikerüli a szabály alapú modellek hibáit
- A top-down és útvonal szintű modell erősségeit kombinálja, hogy kiküszöbölje a gyengeségeket

PATH LEVEL ÉS TOP DOWN MODELLEK INTEGRÁLÁSA

1) Felbecsüli a csatorna avulási arányát és az egyedi csatorna felemelkedést a **Top Down** modellből

2) A Markov-lánc **útvonal szintű** dinamikáját használja fel a csatorna emelkedési és avulási arányaival, a **Top Down** modellből kinyerve

3) Ez a kimenetet átadja az **attribúciós motomak** a napi attribúcióért a tranzakciós útvonalak között, hogy elkészítse az attribúciós súlyozást minden egyes útvonalra, csatornánkénti interakciós pontokkal

Az *egyedi tranzakciós lépés*hez rendelt érték a modell **kimeneti együtthatóján** és a **hozzáadott útvonal specifikus attribútumokon** alapul.

Beleértve az útvonalban elfoglalt pozíciót, konverzió idejét, relatív emelkedését összehasonlítva más csatornákkal az útvonalon, eltekintve attól, hogy kattintásról vagy megjelenésről van szó.

AZ ALGORITMUS ALAPÚ MODELLEK

Markov-láncok tulajdonságai

Először szükséges bevezetnünk az egy-lépéses átmenet-valószínűség:

$$p_{ij} = \Pr(X_1 = j \mid X_0 = i).$$

és az n-lépéses átmenet-valószínűség fogalmát:

$$p_{ij}^{(n)} = \Pr(X_n = j \mid X_0 = i)$$

Előbbi az i állapotból egy lépéssel a j állapotba kerülés, utóbbi az n . lépés után a j állapot valószínűségét fejezi ki, feltéve, hogy $\Pr(X_0 = i) > 0$

Az n-lépéses átmenet-valószínűségekre teljesül a Chapman–Kolmogorov-egyenlőség, amely $0 < k < n$ esetén minden k -ra az alábbi:

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}.$$

Egy Markov-lánc kezdeti eloszlása egy (sor)vektor, mely az alábbi módon értelmezett

$$d(0) = (\Pr(X_0 = x_0), \Pr(X_0 = x_1), \dots, \Pr(X_0 = x_s)).$$

Míg az abszolút eloszlása az n -edik időpillanatban

$$d(n) = (\Pr(X_n = x_0), \Pr(X_n = x_1), \dots, \Pr(X_n = x_s)).$$

Ahol

$$\Pr(X_n = j) = \sum_{r \in S} p_{rj} \Pr(X_{n-1} = r) = \sum_{r \in S} p_{rj}^{(n)} \Pr(X_0 = r). \quad j=1, 2, \dots, s. \text{ s féle állapot van.}$$

Azonban egy Markov-lánc lehet az időtől független, ilyenkor d eloszlása, azaz $d(n)$ nem függ n -től, azaz $d(n) = d(0)$ minden n -re. Ekkor szokás az eloszlást egyszerűen X_n -vel jelölni.

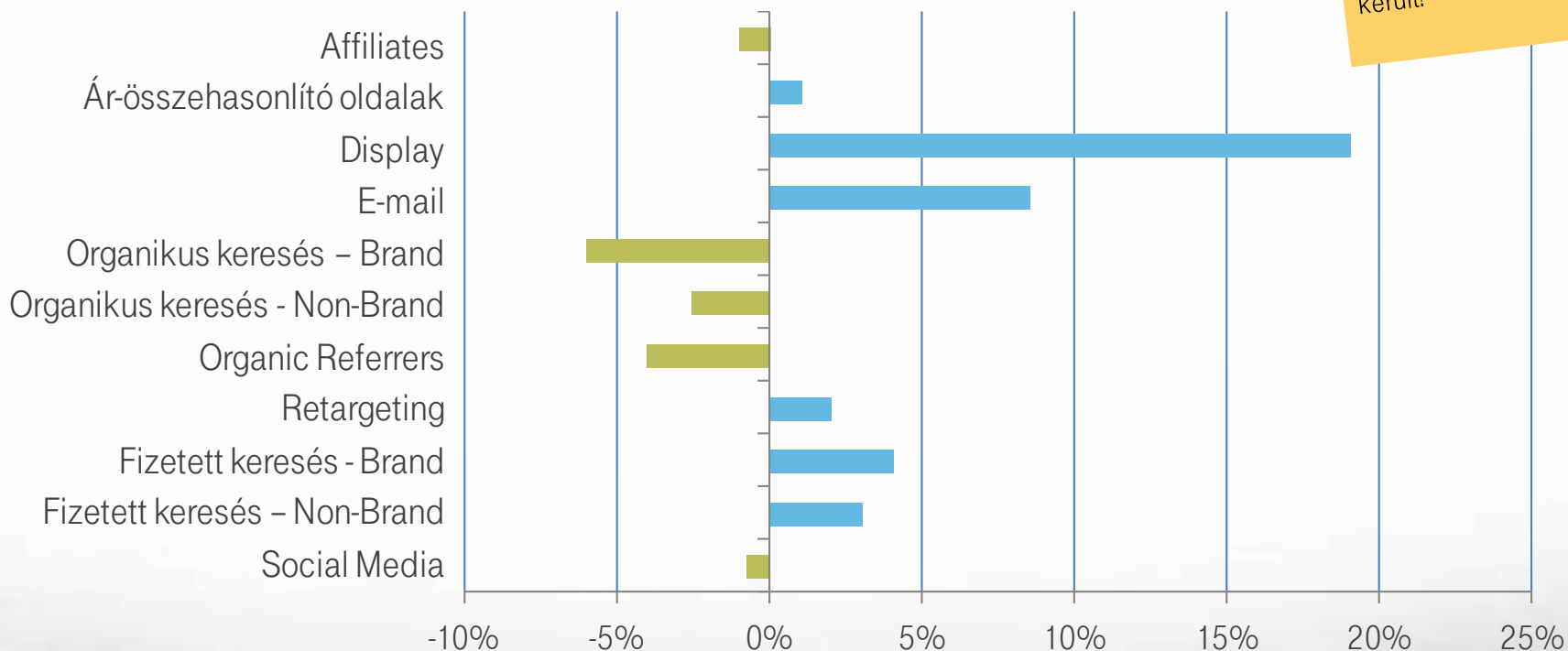
Na ezt nem kell tudni!

AZ ALGORITMUS ALAPÚ MODELLEK

Path és top-down eredményessége

Csatorna szerep változása az Altitude (távolsági) modell használatával

Százalékos változás a többlépcsős csatorna bevételekben, a Last Click modellből továbbfejlődve a Távolsági modellre

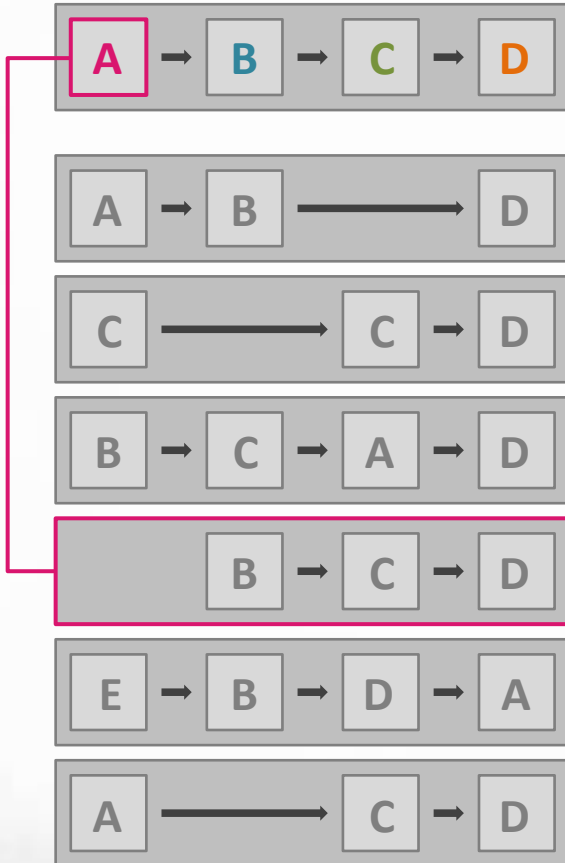


68%-a

A modellben lévő bevételeknek máshová került!

AZ ALGORITMUS ALAPÚ MODELLEK

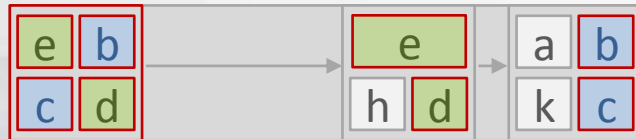
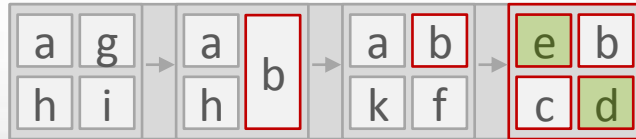
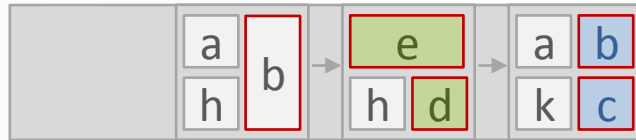
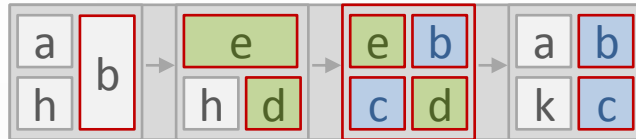
A Path (útvonal) esetleges gyengeségei



- ✗ Megelőző és következő érintkezési pontokon alapul
Nem alkalmazható pontos folyamat tervezésre
- ✗ Adathiányosságokkal és mintavétellel küzd
Az üzleti döntések korlátozott adatokon alapulnak
- ✗ Becslést használ, amely csökkenti a pontosságot
Az ajánlatok pontossága ingadozik

AZ ALGORITMUS ALAPÚ MODELLEK

2. A multi-dimenzionális modell erősségei



Multi-dimenzionális megközelítés



A javaslatok egyszerűbben felhasználhatóak



Nincs adathiányból vagy mintavételből adódó gond



Az üzleti döntések a teljes adatmennyiségen alapulnak



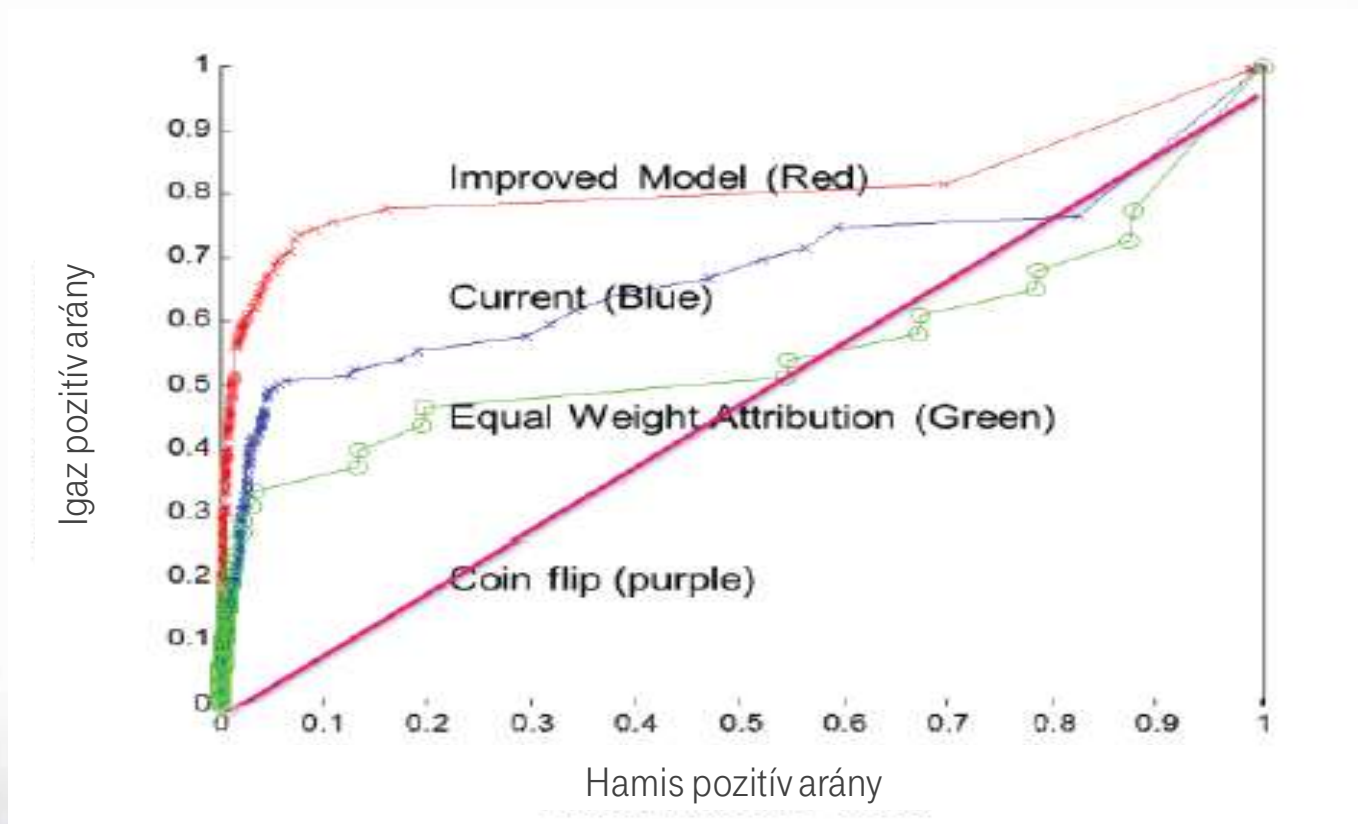
Nincs szükség becslésre, egzakt adatokat használ



Az ajánlatok pontossága minimális ingadozással jár

AZ ALGORITMUS ALAPÚ MODELLEK – ÖSSZEHASONLÍTÁS

A teljesen vak (pénzfeldobás) döntéshez képest az egyenlő eloszlású szabály sikeressége, a kezdeti algoritmus alapú és az öntanulóan fejlődő modell sikeressége.



A MODELLEK

Érettségi lépcsőzetesség

Attribúciós alapok

Átlagos útvonal hossza,
Többlépéses útvonalak,
Csatorna pozíció

Köztes állapot adat alapú értés

Média teljesítmény
jelentés, Modell
összehasonlítás

Attribúciós teljesítmény

Kampány ROI,
Csatorna teljesítmény,
Kulcsszó, Elhelyezés,
Kreatív profitabilitás,
Médiamix
optimalizálás

Továbbfejlesztett viselkedési adatok és stratégia

Csatornák egymásra
hatása és Halo-hatás,
Viselkedés- és
mediatervezés
ügyfélszegmens
alapján

MEGNÖVEKEDETT ATTRIBÚCIÓS TUDÁS ÉS ADOPTÁCIÓ

ÖSSZEFOGLALÓ I.

- Olyan **predikcióra** képes, ami nem a múlt statikus elemzésén alapul (jóslás), hanem mindig **a jelent vetíti ki**, változásokat beépíti (újratervezés)
- **Automatizmust** lehet rá építeni, nem csak a kampány végrehajtásban (pl. licit állítás), hanem az **elemzésben, tudás feltárásban** is
- Közel **valós időben** képes a **modellezést** is és a **végrehajtást** is elvégezni
- Előre definiált helyett talált **mintázaton alapuló szegmentációra** képes sőt, akár **user/persona szintre** lebontható (1-2-1 marketing)
- Az **összes lehetőséget** kikalkulálja, sok változóval, **nem** csak **mintavétellel** dolgozik és a **nem konvertáló utakat** is figyelembe veszi (multi-dimenzió)

ÖSSZEFOGLALÓ II.

- Gépi tanulási képessége van - ha többször fut, többet tud
- Számptalan, akár harmadik feles **adatforrást** is képes beépíteni
- **Figyelembe vesz** szezonalitást, trendeket, környezetet, több képernyőt, stb
- Több modellt és megközelítést képes **ötvözni**
- Kezdetben a költsékek nagyobb részét **átmozgatja**, majd **optimalizál**
- *Érettségi fok kérdése*

MINDEZÉRT BIG DATA KÉPESSÉGEKEN ALAPUL, ÉS
POKOLIAN DRÁGA...
DE MEGTÉRÜLHET



Q & A

KÖSZÖNÖM A FIGYELMET!

NAGY BARNABÁS

nagy.barnabas@telekom.hu