Augmented Analysis

Iryna Kovalchyk

dyvenia

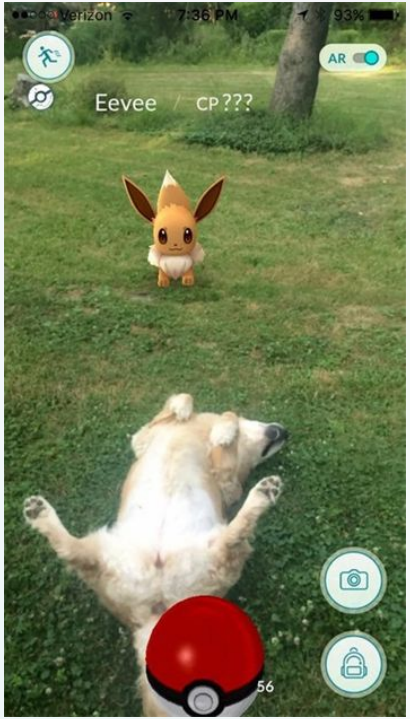
dyvenia

Agenda

- What does Augmented mean?
- 3 Generations of Analytics
- Precise look at AA factors
- How can we implement it?
- Theory on selected algorithms:
  - k-NN,
  - Linear regression,
  - Mean,
  - Mode,
  - Median
- Use case example

dyvenia

What does Augmented mean?



- Augmented analysis (AA): ML and AI
- Augmented reality: combines the real world and computer-generated content
- For example Pokémon GO

dyvenia

3 Generations of Analytics

Traditional analytics	Self-service analytics	Augmented analytics
Driven by IT	Driven by the business	Driven by AI and machine learning
Limited user autonomy*	More user autonomy*	True user autonomy*
Sophisticated tools for data and analytics professionals	User-friendly interface	AI tools and guided processes
Focus on reporting at scale	Focus on user-driven insights	Focus on fast, deep, previously hidden insights

*Autonomy in this case is the ability of user to do analytics without help of other department/specialists e.g. IT-stuff.

dyvenia

Precise look at AA factors

- What makes it possible to focus on fast, deep, previously hidden insights?
  - Using augmentation on different levels, such as data cleaning, data processing/transformation, and visualization, as well as, summarizing patterns
- AI tools and guided processes
  - Numerous frameworks such as TensorFlow, PyTorch, CNTK

dyvenia

How can we implement it?

- Used in banks, financial services, and insurance,
- Promotion of marketing strategies and customer loyalty policies (leads customers to gain trust and invest more),
- Useful in different stages of working with data, since its methodology is based on ML, NLP, and AI.

dyvenia

k-NN

k-nearest neighbors algorithm

The input consists of the k closest training examples in a data set.

The output depends on whether k-NN is used for classification or regression.

dyvenia

Linear regression

An algorithm that provides a linear relationship between an independent variable and a dependent variable to predict the outcome of future events.

Mean, Median, Mode

The **mean** (average) of a data set is found by adding all numbers in the data set and then dividing by the number of values in the set.

The **median** is the middle value when a data set is ordered from least to greatest.

The **mode** is the number that occurs most often in a data set.

dyvenia

Example

Let's check out a few notebooks

Thanks!

Contact us:

ul. Czyżówka 14/0.3,
30-526 Kraków
Poland

ikovalchyk@dyvenia.com
www.dyvenia.com