

AI Identifies Instagram Drug Dealers With Near 95% Accuracy

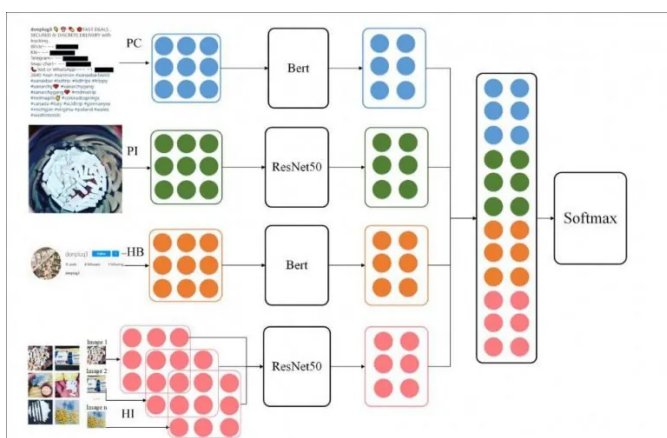
Originally published at <https://www.unite.ai/ai-identifies-instagram-drug-dealers-with-near-95-accuracy/>



Researchers in the US have developed a multimodal machine learning system that's capable of identifying the accounts and posts of drug dealers on Instagram, by analyzing a variety of content, including image content.

The [research](#), entitled *Identifying Illicit Drug Dealers on Instagram with Large-scale Multimodal Data Fusion*, is a collaboration between three researchers at West Virginia University and one from Case Western Reserve University.

To facilitate the project, the researchers created a database called *Identifying Drug Dealers on Instagram* (IDDIG), featuring 4000 user accounts, with 1,400 the accounts of drug dealers, and the remainder as a control group to test the identification process.

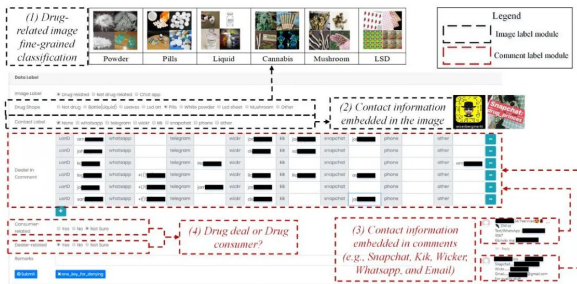


The framework of the multimodal dealer-detection system. The model includes posted images, posted comments, as well as information from homepage images and biography texts posted on the homepage. Source:

<https://arxiv.org/pdf/2108.08301.pdf>

Initial testing of the technique reports almost a 95% accuracy rate in identifying Instagram-based drug dealers, and the framework has also led to a hashtag-based community detection project designed to discover changing signifiers of activity related to the sale of illegal drugs, utilizing geographical factors and identification of specific drug types.

Since the database developed for the project required manual labeling, the framework features a user-friendly annotation system, which uses a classification system based on Google's Bidirectional Encoder Representations from Transformers (BERT), as well as ResNet-based image classification.



The web-based annotation system (with additional indications by the paper's authors) for IDDIG.

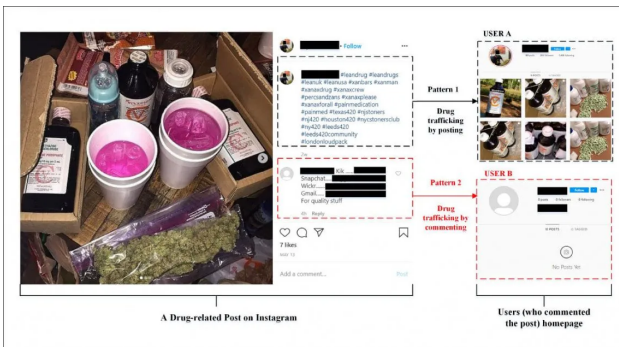
Spotting the Dealers In Drug-Related Conversations

Recreational drugs are discussed in a wide number of contexts across social media platforms such as Instagram. Many of those posting are consumers rather than sellers. Depending on the regulations in their locality, and the possibility of prescription medicine even in localities that differ in their drug legislation, they may also be *legal* consumers.



Drug-related images brought into the project's database.

Additionally, drug dealers' behavior on Instagram is not always explicit; frequently the dealers advertise via comments and hashtags instead of multimedia posts, which would in general be easier to identify as 'drug dealing' content, for both human and machine oversight systems. Therefore hashtags and comment activity have been incorporated as identifying assets in the new system.



Multiple patterns of the drug trade on Instagram posts.

In addition to BERT-based text analysis and ResNet-derived image investigation, the work incorporates feature-level multimodal data fusion, as proposed in the 2016 IEEE [paper](#) *Discriminant Correlation Analysis: Real-Time Feature Level Fusion for Multimodal Biometric Recognition*.

Hashtags as Seeds for a Database

The project's web-scraping mechanism begins its journey to the identification of drug-dealing accounts by tracing the path of 200 drug-related hashtags identified by domain experts, using the [hashtag search](#) API.

Images in posts that use the hashtags are then classified using a VGG-16-based binary [classification model](#). Images that correlate to known drug imagery are then saved in the system, and the post converted to a JSON object for later retrieval.

The framework then extends out to related comments and information (both text and images) contained in the homepage of posters who have participated in the hashtag, and whose content has been flagged as drug-related. In this way 10,000 potential posts and 23,034 user homepages were ingested into the dataset.

Since drug-related hashtags evolve constantly to evade pattern detection and the attention of the authorities, any new hashtags in the flagged post which were not part of the seed hashtag collection are noted and recorded for future use.

After labeling in the web-based interface (see image above), multimodal data fusion has to accommodate the fact that not all posts are going to contain all four possible types of data. Therefore the algorithm is able to tolerate nine out of a total of 16 sub-points among the four data types, using concatenation and fused features, where missing elements will correspond to zero in the calculation.

NetworkX

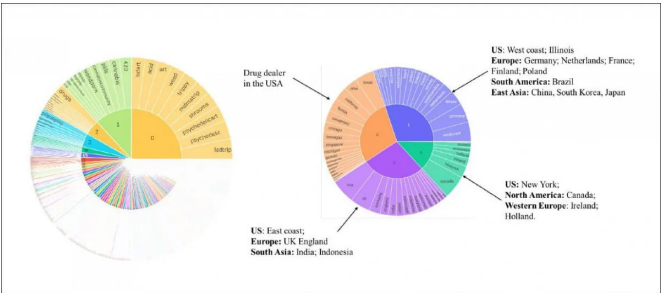
The dataset is finally utilized via the [NetworkX Python language package](#) proposed in 2008 by the Los Alamos National Laboratory at New Mexico. Network X has been used extensively in large-scale operations, including graphs with more than 10 million nodes.

By treating the hashtags in the dataset as if they had been included in one post, it was possible for the researchers to generate an undirected drug-related graph for NetworkX to analyze.

The IDDIG dataset was tested across a variety of protocols including Multi-modal Data Fusion, Multi-source Data Fusion, and Quadruple-based Fusion, and achieved accuracy results of up to 95% in terms of identifying drug-related posts and users, by comparison to human-in-the-loop methods of identification.

Performance		Accuracy	Precision	Recall	F1 score
Post-level fusion	PI + PC	94.95%	93.47%	95.17%	94.31%
Homepage-level fusion	HI + HB	91.29%	88.20%	92.60%	90.35%
Text-based fusion	PC + HB	93.68%	90.44%	95.77%	93.03%
Image-based fusion	PI + HI	92.95%	87.57%	97.89%	92.44%
Quadruple-based fusion (Decision-level fusion)	PI + PC + HI + HB	94.48%	96.92%	90.33%	93.51%
Quadruple-based fusion (Feature-level fusion)	PI + PC + HI + HB	95.55%	96.41%	93.35%	94.86%

It was possible also to generate ‘sunburst plots’ revealing broad indicators for geographic disposition of drug-related activity on Instagram, and other possible future lines of inquiry in similar projects.



Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More