

Image-2-Value: Automated Entity Extraction for Product Information using Random Forest

G. Paavai Anand¹, Vaibhav S. Jain², Jaina Solanki³

Assistant Professor (SG), Department of CSE¹

Undergraduate Student, Department of CSE^{2,3}

SRM Institute of Science and Technology, Vadapalani, Chennai, TN, India

Abstract: *E-commerce and digital marketplaces rely heavily on accurate and complete product information to enhance customer experience and improve product findability. However, the conventional process of manual data entry for key product attributes, such as weight, dimensions, and volume, is labor-intensive, time-consuming, and prone to errors, especially across large datasets. Image2Value addresses these challenges by automating the extraction of product information directly from images, leveraging machine learning techniques with a ResNet50 model for deep feature extraction and a Random Forest classifier for entity prediction. By automatically identifying and classifying attributes from product images, Image2Value significantly improves data efficiency and accuracy, reducing human error while maintaining consistency across various applications in e-commerce, healthcare, and more. This system's scalability allows it to handle diverse product types and large datasets with ease, adapting seamlessly to new domains as needed. Furthermore, the standardized data output produced by Image2Value ensures compatibility with existing platforms, providing a practical, reliable, and innovative approach to enriching digital product information..*

Keywords: Image-2-value , Random Forest Classification, Machine Learning, ResNet50, OCR