

Lorem ipsum dolor sit amet

Loro oooo

Consectetuer adipiscing elit. Aliquam erat volutpat. Integer tempor. Nullam dapibus fermentum ipsum. Quisque porta. Etiam sapien elit, consequat eget, tristique non, venenatis quis, ante. Etiam quis quam. Nullam sapien sem, ornare ac, nonummy non, lobortis a enim. Nullam sit amet magna in magna gravida vehicula. Fusce suscipit libero eget elit. Ut tempus purus at lorem. Maecenas aliquet accumsan leo. Maecenas fermentum, sem in pharetra pellentesque, velit turpis volutpat ante, in pharetra metus odio a lectus. (Albayrak, Aytekin, Kalaycı 2022)

Maecenas fermentum, sem in pharetra pellentesque, velit turpis volutpat ante, in pharetra metus odio a lectus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Mauris dolor felis, sagittis at, luctus sed, aliquam non, tellus. Donec iaculis gravida nulla. Nulla non lectus sed nisl molestie malesuada. Praesent vitae arcu tempor neque lacinia pretium. Fusce suscipit libero eget elit. Cras elementum. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Mauris suscipit, ligula sit amet pharetra semper, nibh ante cursus purus, vel sagittis velit mauris vel metus. Nullam dapibus fermentum ipsum. Integer malesuada. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam sed tellus id magna elementum tincidunt. Etiam sapien elit, consequat eget, tristique non, venenatis quis, ante. Curabitur sagittis hendrerit ante. Aliquam in lorem sit amet leo accumsan lacinia. Aenean placerat. Fusce suscipit libero eget elit. Et harum quidem rerum facilis est et expedita distinctio.

Lorem ipsum dolor sit amet

Consectetuer adipiscing elit. Aliquam erat volutpat. Integer tempor. Nullam dapibus fermentum ipsum. Quisque porta. Etiam sapien elit, consequat eget, tristique non, venenatis quis, ante. Etiam quis quam. Nullam sapien sem, ornare ac, nonummy non, lobortis a enim. Nullam sit amet magna in magna gravida vehicula. Fusce suscipit libero eget elit. Ut tempus purus at lorem. Maecenas aliquet accumsan leo. Maecenas fermentum, sem in pharetra pellentesque, velit turpis volutpat ante, in pharetra metus odio a lectus. (Albayrak, Aytekin, Kalaycı 2022)

Maecenas fermentum, sem in pharetra pellentesque, velit turpis volutpat ante, in pharetra metus odio a lectus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Mauris dolor felis, sagittis at, luctus sed, aliquam non, tellus. Donec iaculis gravida nulla. Nulla non lectus sed nisl molestie malesuada. Praesent vitae arcu tempor neque lacinia pretium. Fusce suscipit libero eget elit. Cras elementum. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Mauris suscipit, ligula sit amet pharetra semper, nibh ante cursus purus, vel sagittis velit mauris vel metus. Nullam dapibus fermentum ipsum. Integer malesuada. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam sed tellus id magna elementum tincidunt. Etiam sapien elit, consequat eget, tristique non, venenatis quis, ante. Curabitur sagittis hendrerit ante.

Aliquam in lorem sit amet leo accumsan lacinia. Aenean placerat. Fusce suscipit libero eget elit. Et harum quidem rerum facilis est et expedita distinctio.

Lorem ipsum dolor sit amet

Consectetur adipiscing elit. Aliquam erat volutpat. Integer tempor. Nullam dapibus fermentum ipsum. Quisque porta. Etiam sapien elit, consequat eget, (Christen 2012) tristique non, venenatis quis, ante. Etiam quis quam. Nullam sapien sem, ornare ac, nonummy non, lobortis a enim. Nullam sit amet magna in magna gravida vehicula. Fusce suscipit libero eget elit. Ut tempus purus at lorem. Maecenas aliquet accumsan leo. Maecenas fermentum, sem in pharetra pellentesque, velit turpis volutpat ante, in pharetra metus odio a lectus. (Albayrak, Aytakin, Kalaycı 2022)

Maecenas fermentum, sem in pharetra pellentesque, velit turpis volutpat ante, in pharetra metus odio a lectus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Mauris dolor felis, sagittis at, luctus sed, aliquam non, tellus. Donec iaculis gravida nulla. Nulla non lectus sed nisl molestie malesuada. Praesent vitae arcu tempor neque lacinia pretium. Fusce suscipit libero eget elit. Cras elementum. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Mauris suscipit, ligula sit amet pharetra semper, nibh ante cursus purus, vel sagittis velit mauris vel metus. Nullam dapibus fermentum ipsum. Integer malesuada. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam sed tellus id magna elementum tincidunt. Etiam sapien elit, consequat eget, tristique non, venenatis quis, ante. Curabitur sagittis hendrerit ante. (McClain 2023) Aliquam in lorem sit amet leo accumsan lacinia. Aenean placerat. Fusce suscipit libero eget elit. Et harum quidem rerum facilis est et expedita distinctio.

Bibliografie

ALBAYRAK, Osman Semih, AYTEKIN, Tevfik a KALAYCI, Tolga Ahmet, 2022. Duplicate product record detection engine for e-commerce platforms. *Expert Systems with Applications*. Online. květen 2022. Vol. 193, p. 116420–116421. DOI 10.1016/j.eswa.2021.116420. [Accessed 28 duben 2024].

Having a clean product catalog and keeping it complying with the standards of the industry is one of the primary concerns of e-commerce companies. Integrating product data from multiple providers confronts the companies with a challenging issue: duplicate product records. Since it is possible to describe a product with a variety of different words, images and attributes, detecting duplicate product records is a difficult task to overcome. In this work, a novel duplicate record detection engine is proposed for an e-commerce company, Hepsiburada. The engine is developed based on a real-world dataset. In order to build a training set we use text similarity and domain-specific distance metrics for generating candidate duplicate product pairs which are then labeled by human experts. We performed extensive feature engineering and state-of-the-art classification models to determine whether any two products are duplicated or not. The experimental results show that our engine is able to detect duplicate product records with high precision and outperforms the accuracy of non-adaptive methodologies.

CHRISTEN, Peter, 2012. *Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection*. Berlin Heidelberg: Springer. Data-centric systems and applications. ISBN 978-3-642-43001-5 978-3-642-31163-5.

MCCLAIN, Bonny P., 2023. *Python for geospatial data analysis: theory, tools, and practice for location intelligence*. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly.
ISBN 978-1-09-810479-5.