

Avoid Plagiarism (from Writer Inst in several Universities and my editing)

Whats Plagiarsm

- Plagiarism means using another's work without giving credit. If you use others' words, you must put them in quotation marks and cite your source. You must also give citations when using others' ideas, even if you have paraphrased those ideas in your own words (Un of Calf)
- Plagiarism is the practice, whether intentional or not, of using someone else's words or ideas and presenting them as your own (WWash Un)

- AWARE of ONLINE Source
- Internet often support to make PLAGIARISM (with or without intention)
 - Make sure always refers to primary source, and secondary source as supporting
 - If get source from Internet, suggestion trace it till find primary source

Common Type (UNSW)

- Downloading an assignment from an online source and submitting it as your own work.
- Buying, stealing or borrowing an assignment and submitting it as your own work.
- Copying, cutting and pasting text from an electronic source and submitting it as your own work.
- Using the words of someone else and presenting them as your own

- Using the words of someone else and presenting them as your own
- Using significant ideas from someone else and presenting them as your own
- Copying the written expressions of someone else without proper acknowledgment
- Relying too much on other people's material

Cont. (MTR Un Ca)

- Copying or paraphrasing from books or other sources without citing it properly.
- Copying work from another student.
- Working as a group on projects where the instructor requires individual work.
- Buying or copying entire papers or projects done by others.
- Altering information or data.
- Using misleading references.
- Resubmitting previously evaluated work

Spec in Writing

(Miguel R, St.Jhon Un)

- Plagiarism of ideas

Appropriating an idea (e.g., an explanation, a theory, a conclusion, a hypothesis, a metaphor) in whole or in part, or with superficial modifications without giving credit to its originator

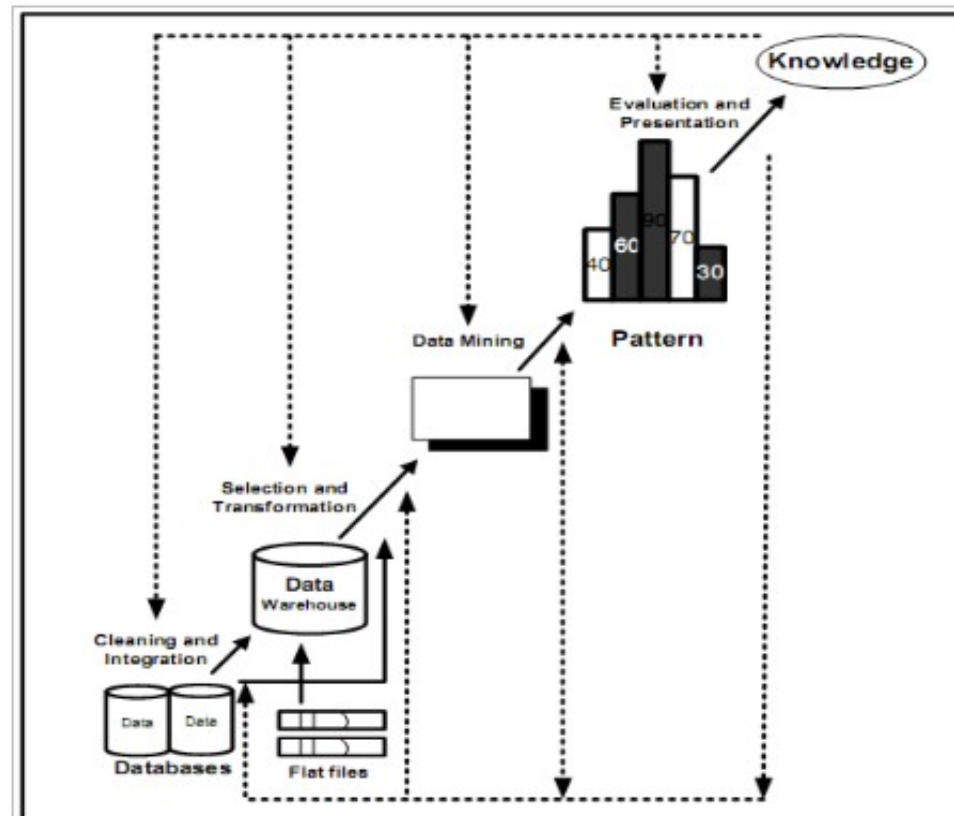
- Use prev idea
- Good ex:

Berdasarkan [3] yang menggunakan modifikasi *Cosine Similarity* dengan konsep *Bag of Words Model*, yaitu menggunakan perhitungan kemunculan *term* pada suatu kalimat dikalikan dengan nilai *idf term* tersebut, maka kemiripan kalimat *s* dapat dihitung menggunakan formula *Cosine Similarity (IDF Modified)* :

$idf_modified_Cosine(x, y) = \frac{\sum_{w \in x, y} tf_{w,x} * tf_{w,y} * (idf_w)^2}{\sqrt{\sum_{x_i \in x} (tf_{x_i,x} * idf_{x_i})^2} * \sqrt{\sum_{y_j \in y} (tf_{y_j,y} * idf_{y_j})^2}}$	(4.3)
---	-------

Sebagai suatu rangkaian proses, data *mining* dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar 1. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan knowledge base.

- Bad one



Gambar 2. Tahap-tahap *data mining*

- Plagiarism of text
 - Copying a portion of text from another source without giving credit to its author and without enclosing the borrowed text in quotation marks
 - Copying a portion of text from one or more sources, inserting and/or deleting some of the words, or substituting some words with synonyms, but never giving credit to its author nor enclosing the verbatim material in quotation marks

Despite the overwhelming poverty, black residents in inner-city neighborhoods verbally reinforce, rather than undermine, the basic American values pertaining to individual initiative. For example, the large survey of the UPFLS found that nearly all the black respondents felt that plain hard work is either very important or somewhat important for getting ahead...Nonetheless, given constraints and limited opportunities facing people in inner-city neighborhoods, it is altogether reasonable to assume that many of those who subscribe to these values will, in the final analysis, find it difficult to live up to them.

(W.J. Wilson, *When Work Disappears*, 1996, p. 67)

PLAGIARISM:

Black residents in inner-city neighborhoods verbally reinforce, rather than undermine, the basic American values pertaining to individual initiative.

OK:

Wilson (1996: 67) argues that "[b]lack residents in inner-city neighborhoods verbally reinforce, rather than undermine, the basic American values pertaining to individual initiative."

- Inappropriate paraphrasing or summarization

Taking portions of text from one or more sources, crediting the author/s, but only changing one or two words or simply rearranging the order, voice (i.e., active vs. passive) and/or tense of the sentences

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk *data mining* tidak hanya berasal dari satu *database* tetapi juga berasal dari beberapa *database* atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

Self-plagiarism

- Redundant and Duplicate Publications

Instances in which dual publication may be acceptable

- Salami Slicing (i.e., data fragmentation) or data augmentation

In order to reach specific result, than do fragmentation or augmentation

Guide to Avoid Plagiarsm

- An ethical writer **ALWAYS** acknowledges the contributions of others and the source of his/her ideas
- Any verbatim text taken from another author must **be enclosed in quotation marks**
- We must **always acknowledge every source** that we use in our writing; whether we paraphrase it, summarize it, or enclose it quotations

- When we **summarize, we condense**, in our own words, a substantial amount of material into a short paragraph or perhaps even into a sentence
- Whether we are paraphrasing or summarizing we must **always identify the source** of our information
- When paraphrasing and/or summarizing others' work we must **reproduce the exact meaning** of the other author's ideas or facts using our words and sentence structure.

- In order to make substantial modifications to the original text that result in a proper paraphrase, the author **must have a thorough understanding of the ideas and terminology** being used
- A responsible writer has an ethical **responsibility** to readers, and to the author/s from whom s/he is borrowing, to **respect** others' ideas and words, to credit those from whom we borrow, and whenever possible, to use one's own words when paraphrasing

- **When in doubt** as to whether a concept or fact is common knowledge, provide a citation
- Authors who submit a manuscript for publication containing data, reviews, conclusions, etc., that have already been disseminated in some significant manner (e.g., published as an article in another journal, presented at a conference, posted on the internet) **must clearly indicate** to the editors and readers the nature of the previous dissemination

- If the results of a **single complex study** are best presented as a 'cohesive' single whole, they **should not be partitioned into individual papers**. Furthermore, if there is any doubt as to whether a paper submitted for publication represents fragmented data, authors should enclose other papers (published or unpublished) that might be part of the paper under consideration (Kassirer & Angell, 1995)

- Because some instances of plagiarism, self-plagiarism, and even some writing practices that might otherwise be acceptable (e.g., extensive paraphrasing or quoting of key elements of a book) can constitute copyright infringement, authors are **strongly encouraged to become familiar with basic elements of copyright law**

- While there are some situations where text recycling is an acceptable practice, it may not be so in other situations. Authors are urged to adhere to the spirit of ethical writing and **avoid reusing their own previously published text**, unless it is done in a manner consistent with standard scholarly conventions (e.g., by using of quotations and proper paraphrasing)

- Authors are **strongly urged to double-check their citations.** Specifically, authors should always ensure that each reference notation appearing in the body of the manuscript corresponds to the correct citation listed in the reference section and that each source listed in the reference section has been cited at some point in the manuscript. In addition, authors should also ensure that all elements of a citation (e.g., spelling of authors' names, volume number of journal, pagination) are **derived directly from the original paper, rather than from a citation that appears on a secondary source.** Finally, authors should ensure that credit is given to those authors who first reported the phenomenon being studied

- The references used in a paper should only be those that are **directly related to its contents**. The intentional inclusion of references of questionable relevance for purposes of manipulating a journal's or a paper's impact factor or a paper's chances of acceptance is an unacceptable practice

- Authors should follow a simple rule: Strive to obtain the **actual published paper**. When the published paper cannot be obtained, cite the specific version of the material being used, whether it is conference presentation, abstract, or an unpublished manuscript

- Generally, when describing others' work, **do not rely on a secondary summary of that work**. It is a deceptive practice, reflects poor scholarly standards, and can lead to a flawed description of the work described
- If an author must rely on a secondary source (e.g., textbook) to describe the contents of a primary source (e.g., an empirical journal article), s/he should **consult writing manuals** used in her discipline to follow the proper convention to do so. Above all, always indicate the actual source of the information being reported

- When borrowing heavily from a source, authors should always craft their writing in a way that makes clear to readers which ideas are their own and which are derived from the source being consulted.
- When appropriate, authors have an ethical responsibility to report evidence that runs contrary to their point of view. In addition, evidence that we use in support of our position must be methodologically sound. When citing supporting studies that suffer from methodological, statistical, or other types of shortcomings, such flaws must be pointed out to the reader

- Authors have an ethical obligation to **report all aspects** of the study that may impact the independent replicability of their research.
- Researchers have an ethical responsibility to **report the results of their studies according to their a priori plans**. Any post hoc manipulations that may alter the results initially obtained, such as the elimination of outliers or the use of alternative statistical techniques, must be clearly described along with an acceptable rationale for using such techniques.

- Authorship determination should be discussed prior to commencing a research collaboration and should be based on established guidelines, such as IEEE, InderScience.
- Only those individuals who have made substantive contributions to a project merit authorship in a paper

- Faculty-student collaborations should follow the same criteria to establish authorship. Mentors must exercise great care to neither award authorship to students whose contributions do not merit it, nor to deny authorship and due credit to the work of students.
- Academic or professional ghost authorship in the sciences is ethically unacceptable.

- Authors must become **aware of possible conflicts** of interest in their own research and to make every effort to disclose those situations (e.g., stock ownership, consulting agreements to the sponsoring organization) that may pose actual or potential conflicts of interest.”