Methods to differentiate between machine written text and human written text

Research Studies
-Based on a study these are some features used to distinguish between original human generated text and AI generated text

 Asked chat GPT to rephrase original human generated text
Observation
-chatGPT describes experiences as opposed to the feelings
-avoids the use of personal pronouns
-Does not use aggressive or rude language
- text lacked detailed found in original text
- writing style observed to be extremely polite
- cannot produce responses that include metaphors, irony, or sarcasm
- certain words were overly repeated ie. the, it , it
- human generated text contained many typos

AI generated text tends to be

-**Uncreative:** texts tend to be uncreative or lack originality, writing styles that were unique and expressive chat GPT was not able to emulate a similar writing style
- **Repetition and Specificity:** AI generated text tends to repeat key phrases and sentences, this reputation style is unlikely to occur in human text. Trained on a large corpus of text, meaning that it may repeat certain phrases or expressions more often than a human writer would.
 -AI text contains overly specific details, such as excessively precise dates or locations,
- **Errors or typos:** both human and AI text have errors however human errors tend to be more spelling or grammatical errors. AI text tend to have more systematic errors such as repeating words or using the wrong word
- **Lacks Personal experience:** AI generated text is based on patterns from data and therefore does not include any personal anecdote or personal stories
-**Cohesion**: Human generated text has a clear flow and thoughts follow a logical manner whereas AI generated text is often poorly organized with no clear flow
Sentiment:AI text lacks human emotion and tone, output may lack emotional depth
Consistency: Ai text is based on data patterns and often times there is randomness in the output text
-**Lack of Contextual Awareness:** output may not always be in line with current events or culturally relevant information because it's unaware of current world events
-AI-text detection are currently best suited to benefit English speakers, there's a limited efficacy outside English
-Reports facts bluntly, no additional commentary or individuality

Resources:
https://www.zdnet.com/article/chatglypt-or-human-text-heres-one-surprising-way-to-tell-the-difference/
https://medium.com/swlh/detecting-ai-generated-text-b56b3e8ca730
https://towardsdatascience.com/telling-apart-ai-and-humans-3-text-and-humor-c13e345f4629

https://www.angmohdan.com/chatgpt-vs-human-generated-text-how-to-spot-the-difference/#:~:text=Lack%20of%20creativity%3A%20Humans%20can,not%20be%20able%20to%20replicate.

https://medium.com/youve-been-informed/can-you-tell-the-difference-between-ai-and-human-authorship-444d41047f7f

Using NPL
-NPL uses features like keyword extraction and topic classification to understand intent of language of text and generate a response

-Large language models such s Open AI, uses an AI classifier
Work best when there is a large amount of text available

NPL techniques:

Syntactic Analysis- analyzes text based on basic grammar rules, this is done through Tokenization,Stop-word removal,and Lemmatization & stemming

Semantic Analysis- captures the meaning of text, this is done through Word sense disambiguation and Relationship extraction

https://monkeylearn.com/blog/nlp-ai/
https://enterprisersproject.com/article/2020/2/artificial-intelligence-ai-vs-natural-language-processing-nlp-differences