A Survey Summary of Automatic Sarcasm Detection

PROBLEM STATEMENT: Sarcasm Detection is the Task of identifying Irony containing Utterances in Sentiment Bearing Text. However, the figurative and creative nature of sarcasm poses a great challenge for affective computing systems performing sentiment analysis. Since Sarcasm enables one speaker or writer to conceal their true intention of contempt and negativity under a guise of overt positive representation. Thus, recognizing Sarcasm and verbal irony is critical for understanding people's actual sentiments and beliefs.

Examples for SARCASM Statements: The figurativeness and subtlety inherent in its sentiment display, a positive surface with a contemptuous intent (e.g., "He has the best taste in music!"), or a negative surface with an admiring tone (e.g., "She always makes dry jokes!"), makes the task of its identification a challenge for both humans and machines.
Empirical studies of this linguistic device refer to methods to predict if a given user-generated text is sarcastic or not. From a computational perspective, this task is formulated as a binary classification problem.

For Sarcasm, its identification is reliant on common sense and connotative knowledge that come naturally to most humans but makes machines struggle when extra textual information is essentially required. Sarcastic utterances are often expressed in such nuanced ways that should be distinguished from a similar phenomenon called humble - bragging, which is a self-representational verbal strategy that appears as a complaint concealed within a bragging as in "I am a perfectionist at times, it is so hard to deal with".

TECHNIQUES for AUTOMATIC SARCASM DETECTION: There are 2 types of Detection Methods:

a) **Content-Based Methods:** Models investigated in this section base their identification of sarcasm on lexical and pragmatic indicators in English language use on social media. There are 3 different Approaches considered within:

Rule-Based Approach: Rule-based attempts look for evidence and indicators of sarcasm and rely on those in forms of rules. Hashtags (or their equivalent, given the social media platform) have been utilized by users to de note sarcasm on Twitter (e.g., #sarcasm, #not) or on Reddit. Another approach is to consider a parsing algorithm that looks for sentiment-bearing situations and identifies sarcasm in forms of a contradiction of negative (or positive) sentiment and positive (or negative) situation.
They also look for the co-occurrence of interjection hyperbolic words like "wow", "yay", etc. at the start of tweets, and intensifiers like "absolutely", "huge" e.g., "Wow, that's a huge discount, I'm not buying anything!! #sarcasm."

Feature Set Approach: Here, we consider the salient textual features effectively utilized toward the detection of sarcasm. Most studies use bag-of-words to an extent. Nonetheless, in addition to these, the use of several other sets of features have been reported.
Another approach is to consider set of humour dependent or irony-dependent features related to ambiguity, unexpectedness, and emotional scenario. Ambiguity features cover structural, semantic ambiguity, while unexpectedness features gauge semantic relatedness.
In addition to a Rule Based classifier, use a set of patterns, specifically positive verbs and negative situation phrases, as features.

Several Features can be used for this Approach like: Length of tweet, punctuation marks, emoticons, Cognitive features extracted from eye-movement patterns of human readers , etc.

Learning-Based Methods: Three Methods used here – Supervised Learning, Semi-Supervised Learning and Unsupervised Learning.

- **Supervised Learning:** Here, most work on statistical detection of sarcasm has relied on various combinatory forms of Random Forests (RF), Support Vector Machines (SVM), Decision trees (DT), Naive Bayes (NB) and Neural Networks (NN).

While rule-based approaches mostly rely upon lexical information and require no training, machine learning invariably makes use of training data and exploits different types of information sources (or features), such as bags of words, syntactic patterns, sentiment information or semantic relatedness.

- **Semi-Supervised Learning**: This form of machine learning, which falls between unsupervised learning and supervised learning, uses a minimal quantity of annotated (labelled) data and a large amount of un-annotated (unlabelled) data during training. The presence of the unlabelled datasets and the open access to the unlabelled datasets is the feature that differentiates the semi-supervised from supervised learning.

- **Unsupervised Learning:** Mostly focussed towards Research, most approaches here are clustering-based, which are mostly applicable to pattern recognition. They build on probabilistic topic models originally defined for sentiment analysis. Examples for Unsupervised Algorithms used: Fuzzy C Means (FCM) Clustering.

b) **Context-Based Models**: Making sense of sarcastic expressions is heavily reliant on the background knowledge and contextual dependencies that are formally diverse. Example - "I'm sure Hillary would've done that, lmao." requires prior knowledge about the event, like familiarly with Hillary Clinton's perceived habitual behaviour at the time the post was made. Means the proposed Models need both Content and Contextual Information required for Sarcasm Detection. Also, including features independent of the text leads to ameliorating the performance of sarcasm models. To this end, studies take three forms of context as feature: **1)** Author Context **2)** Conversational Context and **3)** Topical Context.

Another approach is to consider various user embedding techniques that encode users' stylometric and personality features to improve their sarcasm detection models. When used along with content-based feature extractors such as Convolutional Neural Networks (CNNs), significant boost in the classification performance is achieved.

Another approach can be as a sequence classification problem by leveraging the natural shifts in various emotions over the course of a piece of text. To propose a semi-supervised method for contextual sarcasm detection in online discussion forums. They adopt author and topic sarcastic prior preference as context embedding that provides a simple yet representative background knowledge.

Sarcasm Detection using Pre-trained Language Models: To capture figurative language phenomena and the difficulties of data annotation, transfer learning approaches are gaining attention in various domain adaptation problems. In particular, the utilization of pre-trained embeddings such as Global Vectors (GloVe) and ELMo or leveraging Transformer seq2seq methods such as BERT (Bidirectional Encoder Representations from Transformers), RoBERTa and XLNet are witnessing a surge.

DATASETS (which can be used for Studies on Sarcasm Detection): The Datasets used for computational studies on Sarcasm Detection are divided into Three Categories - Short Text (e.g., Tweets, Reddit), Long Text (e.g., discussions on forums), Transcripts (e.g., conversational transcripts of a TV show or a Call Centre).
Short Text (the dominant form of expression on social media, mostly as a direct result of restriction on text length.) can contain only one (possibly sarcastic) utterance, whereas long text may contain a sarcastic sentence among other non-sarcastic sentences that could potentially function as context.

CONCLUSIONS: Sarcasm Detection research has seen a significant surge in interest in the past few years, which justifiably calls for an investigation. Three major approaches in Sarcasm Detection research: 1) Use of Hashtag-Driven supervised learning toward building Annotated Datasets 2) Semi-Supervised Pattern Extraction to identify Implicit Sentiment. 3) The Utilization of Extra-Textual Information as Context (e.g., user's characteristic profiling).