Update:

1.   Unique mapping of regulations extracted from case files and regulations possible
    2. Not possible to directly match condition-action statements extracted from regulation docs with information in case files because -

    I. Highly context dependent on case
    II. Can't match all violations or statements to a specific condition or action
    III. Information summarized in evidence might be added as a table which is distorted after parsing
    IV. Context for a particular violation spans over paragraphs.

    Example:

Regarding consideration amount for off market trades Mr. Mahesh Thanvi stated that he received two cheques each of Rs.20.00 lacs from Shri Bharat Shah towards the consideration amount for the off market trades of 1,50,00,000 shares. Mr. Mahesh
Thanvi deposited one cheque in his personal account in Bank of Rajashthan, Andheri Branch, Mumbai on 5th January, 2005 and one cheque of Rs.20.00 lacs in his personal account in Bank of Baroda, Shivaji Park Branch, Mumbai. Mr. Mahesh Thanvi had given the Bank statements of account with Bank of Rajasthan but did not submit the bank statement of account with Bank of Baroda. From the statement obtained from Bank of Baroda it was observed that there was no deposit of Rs.20.00 lacs. Mr. Mahesh Thanvi denied relationship with any of the off-market transferees and stated that Shri Bharat Shah was introduced to Mr. Mahesh Thanvi by a Sub-Broker Deepak Shah, Mumbai and the buyer (Bharat Shah) advised Mr. Mahesh Thanvi to transfer to different accounts and
Mr. Mahesh Thanvi transferred the shares in different accounts. ==The bank statements verified during investigation do not match the payment details== as submitted by Mr. Mahesh Thanvi in his statements.

3. Found a few related research work that might help in sentence level or phrase level content analysis of case files.

Steps Ahead:

1. Build on some techniques used in related work to extract relevant information