Phase 3

Phase 3

- **Phase 3** is a continuation of **Phase 1 and 2**.
- You use Phase 1 and 2 documents as a ground for extension, and Phase 3 will be going deeper into what has been developed before.
- You extend it by adding more content in all the previous Phases section and by adding a couple of new ones, if you have not had them already:
  - Threats to Validity
  - Related Work
  - Conclusion

Threats to Validity

- It is important to consider the experiment validity already in the planning phase
- The results must be valid for the population of interest If the population is one organization, results have to be valid for this organization
- The four main types of validity threats
  - Conclusion
  - Internal
  - Construct
  - External

Threats to Conclusion Validity

- Concerns with the relationship between treatment and outcome. It affects the ability to draw correct conclusion about relations between treatment and outcome
- Examples
  - Choice of statistical tests
  - Choice of sample size
  - Measurement of the experiment

Threats to Internal Validity

- It can affect the independent variable with respect to causality. The results may indicate a causal relationship, although there is none
- Factors to consider
  - How the models are chosen?
  - Are we choosing the right metrics, features, to learn from?

Threats to Construct Validity

- It refers to the extent to which the experiment setting reflects the theory Factors to consider
  - Are students representative of professionals??
  - Is the dataset reflective of real world scenarios?
  - Are we choosing the right evaluation indicators?

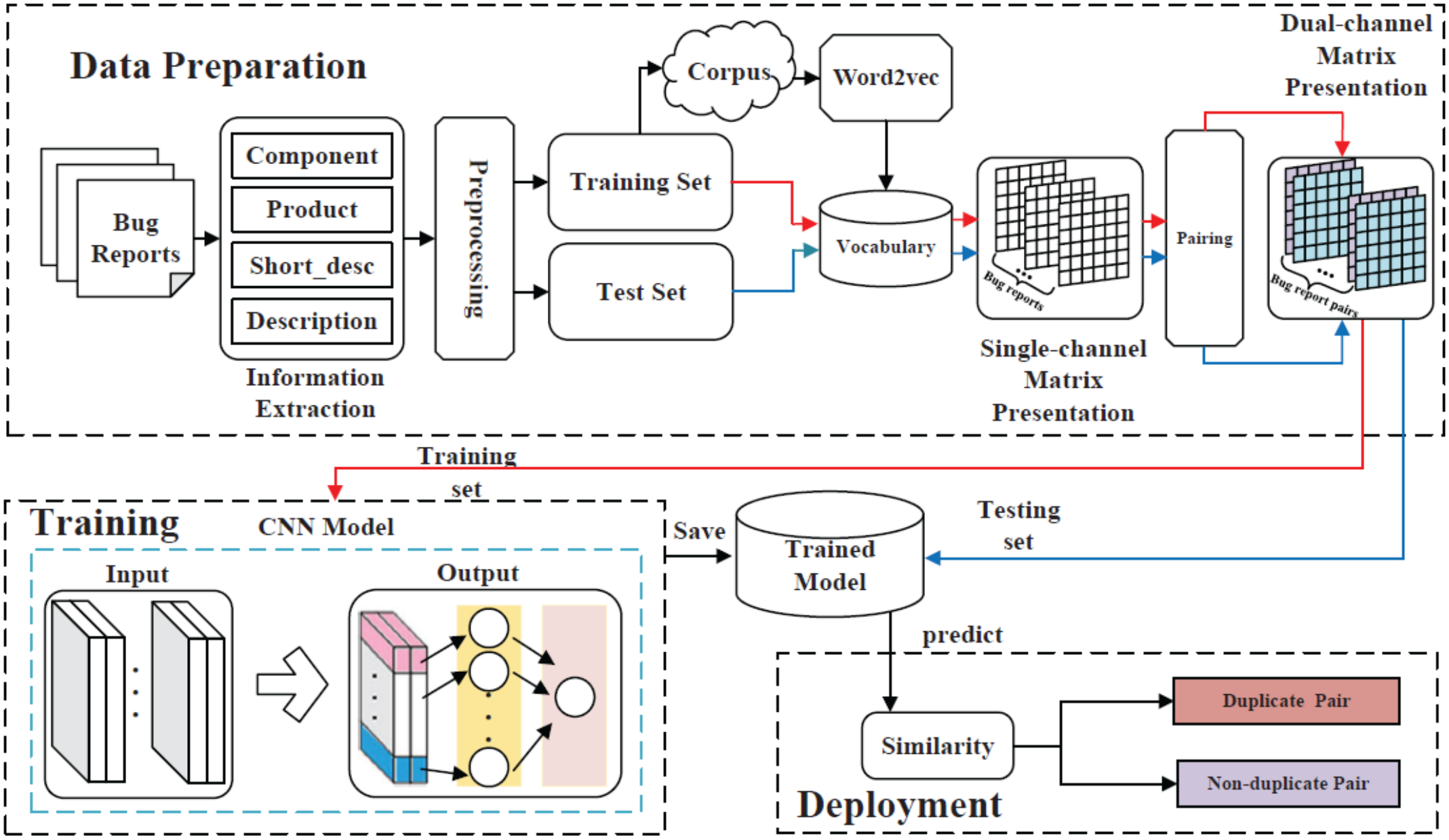
Threats to External Validity

- It limits the ability to generalize the results beyond the experiment setting. These treats are reduced by making the experimental environment more realistic
  - Are we using diverse datasets?
  - Are we using various projects with different characteristics that are close to reality?
  - Wrong environments (too artificial)?

Extending Study Design

- In Phase 2, you were asked to come up with study overview in a black box fashion. Now you need to add one or many figures representing the “white box” version(s).
- Note: If you are using multiple models, you need only to white-box the best performing one.
- White boxing means explaining, in figures, and then text, how exactly the input is **pre-processed**, and/or how the features are being **calculated**, and/or the model is exactly **learning**.

Example of Black-Box overview



Example of its White-Boxing

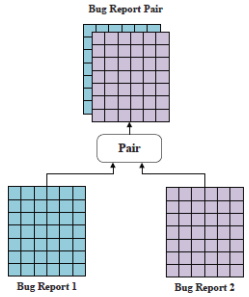


Figure 2: Combining Two Bug Reports into a Bug Report Pair

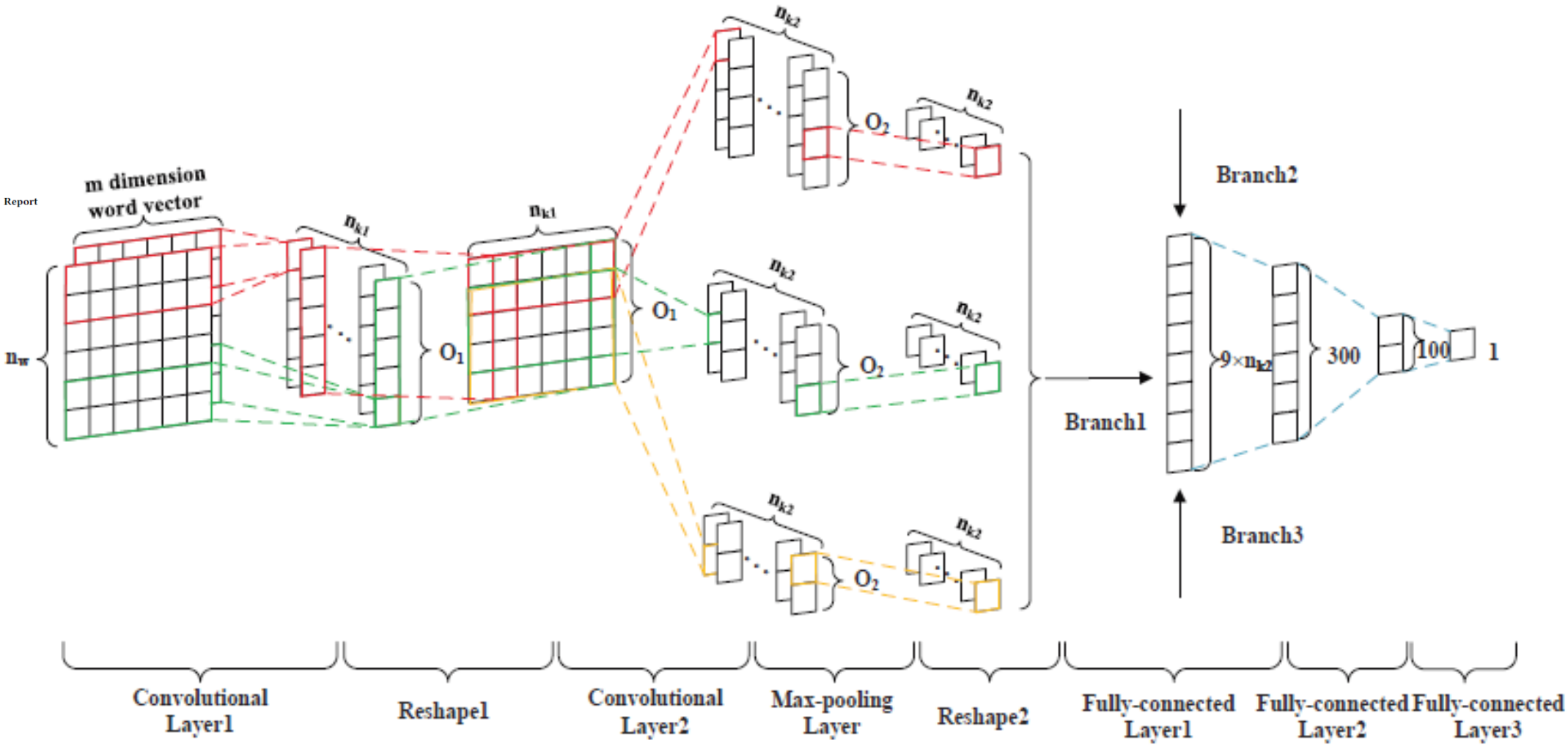


Figure 3: The Overall Workflow of CNN Model

Another Example of Black-Box overview

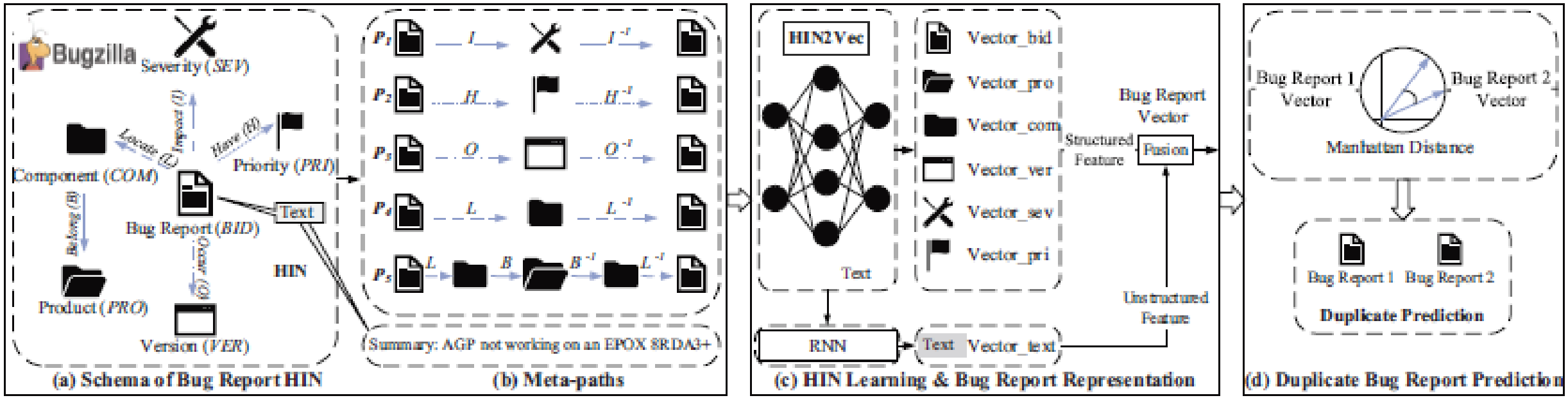


Fig. 1. Overview of HINDBR.

And here is its White-Boxing

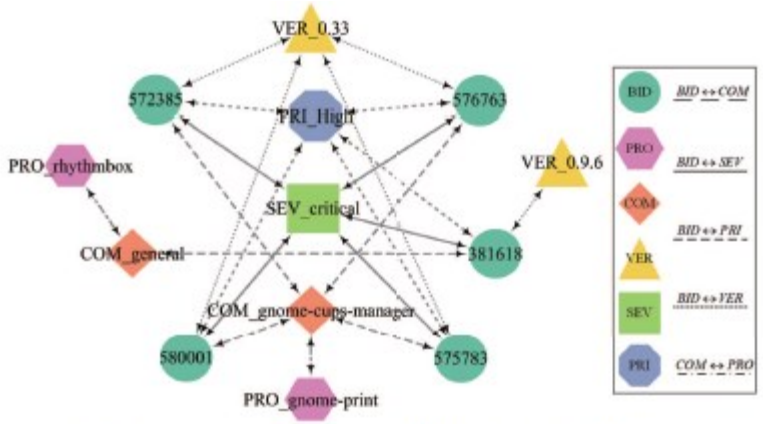


Fig. 2. An excerpt of a bug report HIN for the GNOME project.

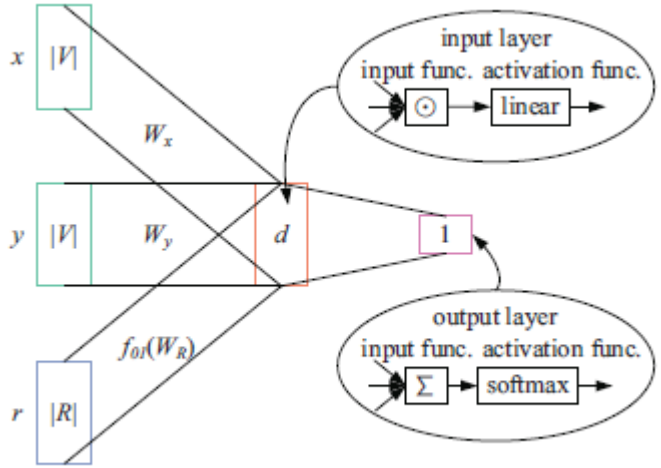


Fig. 4. The HIN2Vec neural network model.

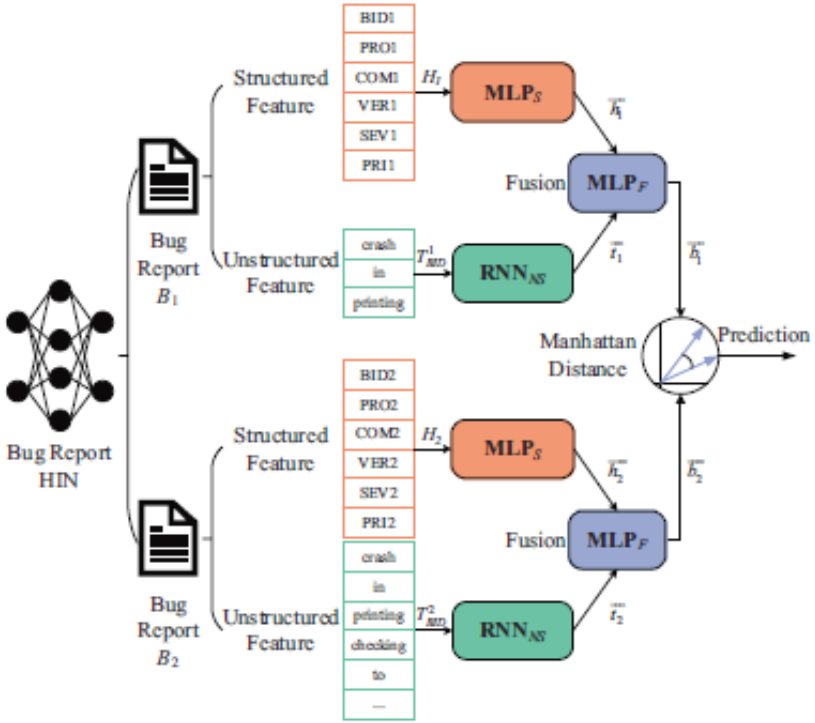


Fig. 5. Detailed structure of HINDBR.

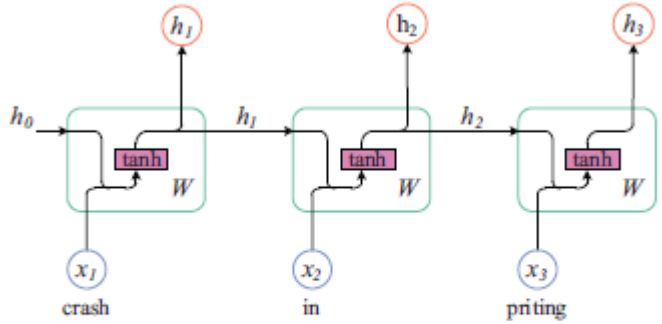


Fig. 6. Sequence embedding using an RNN.

HINDBR embeds the sequence of split word tokens using an RNN (denoted as $\mathbf{RNN}_{NS}$):

Extension of Experiments

- This section now should contain clear interpretation of the results, and not just reporting the numbers.
- You should use the **confusion matrix** to develop conclusions of the performance of your model, and explain why it is successful in some cases and why it is not successful in other cases.
- You should add examples from the confusion matrix, of some **false positives** (and explain why your model made the mistake of selecting), and some **false negatives** (and explain why your model missed it).