Class Notes

Lack of implementation plan

Assessment

1. Validity

- concept validity

- content validity – hidden effects

- correlation validy

2. Accuracy

- false positive and negatives

3. Reliability

- do we get consistent result

4. Comprehensive

- can it handle all possible inputs, even the wrong

Risks and concerts

1. Data subversion

2. Performance degradation in workflows. Acceptance testing by end user

4. Handle attacks – handle deliberate use of bad data

5. Define what are acceptable rates of error.

6. Lack of mitigation for failure

Issues

1. Legal liability and accountability
2. Should and AI have a security clearance?
3. Models and data should be locked down

Other people's AI.

1. Spoofing – generating bogus credentials or biometrics
2. Generation of realistic by fake tax data or financial