

Interprétation de quelques exemples

Ce rapport traite quelques exemples pour l'interprétation de paires de phrases. Plusieurs tests, présentés en tant que " cas ", ont été effectués pour comprendre les choix du modèle.

Remarque : les mots en rouge sont les mots qui ont été changés.

Exemple 1 :

label : inférence

P = " Male in a blue jacket decides to lay in the grass . "

H = " The guy wearing a blue jacket is laying on the green grass "

Explication humaine :

Les mots importants à associer sont " male " et " guy " qui veulent dire la même chose, " blue jacket " dans les deux phrases, " lay " et " laying ", " grass " dans les deux phrases.

Tout est identique (soit on relie, dans l'hypothèse et la prémisse, le même mot ; soit il y a deux mots différents mais qui ont exactement le même sens, comme " male " et " guy ").

Interprétation du modèle :

Le modèle prédit bien une inférence.

- Cas 1 : P = "male" ; H = "guy"

Le modèle prédit bien une inférence.

- Cas 2 : P = "lay" ; H = "laying"

Le modèle prédit bien une inférence.

- Cas 3 : P = "the grass" ; H = "the green grass"

Le modèle prédit bien une inférence.

- Cas 4 : P = "**female** in a blue jacket decides to lay in the grass . " ; H = " The guy wearing a blue jacket is laying on the green grass " (hypothèse inchangée)

Le modèle prédit bien une contradiction : il a vu que " female " est différent de " guy " donc il prend bien en compte la personne dont on parle.

- Cas 5 : P = " Male in a blue jacket decides to lay in the grass . " (prémisse inchangée) ; H = "The guy wearing a **green** jacket is laying on the green grass"

Le modèle prédit une inférence au lieu d'une contradiction. Pourtant le modèle prédit une contradiction pour P = "blue jacket" et H = "green jacket".

On va tenter de changer la couleur, car " green " est peut-être associé à " grass ".

- Cas 6 : P = " Male in a blue jacket decides to lay in the grass . " (prémisse inchangée) ; H = "The guy wearing a **red** jacket is laying on the green grass"

Le modèle prédit une contradiction !

Le modèle voit plusieurs fois le mot green, qui a été associé à "grass" (cas 3) en tant qu'inférence, donc dans le cas 5 il prédit une inférence car beaucoup de relation "inférence" ont été associés.

On peut conclure qu'ici le modèle ne relie pas vraiment les bons mots entre eux.

Exemple 2 :

label : contradiction

P = " A person wearing a straw hat , standing outside working a steel apparatus with a pile of coconuts on the ground . "

H = " A person is burning a straw hat . "

Explication humaine :

Les mots importants à associer sont : "person" dans les deux phrases, "wearing" et "burning" qui ne veulent pas dire la même chose, "straw hat" dans les deux phrases.

"person" et "straw hat" prouve que l'on parle bien de la même personne, tandis que "burning" vient donner la contradiction (il est peu probable que la personne puisse porter le chapeau et le brûler en même temps).

Il faut ignorer la seconde partie de la prémisse car il n'y a pas d'informations sur le chapeau.

Interprétation du modèle :

Le modèle prédit bien une contradiction.

On va voir si le modèle ignore la seconde partie de la prémisse :

- Cas 1 : P = "A person wearing a straw hat ." ; H = " A person is burning a straw hat . " (hypothèse inchangée)

Le modèle prédit bien une contradiction : comme nous, le modèle ignore la seconde partie de la prémisse.

- Cas 2 : P = "wearing" ; H = "burning"

Le modèle prédit bien une contradiction.

- Cas 3 : P = "A person wearing a straw hat ." ; H = "A **dog** is wearing a straw hat ."

Le modèle prédit bien une contradiction. Le modèle comprend que l'on ne parle pas de la même entité (person / dog).

- Cas 4 : P = "A person wearing a straw hat ." ; H = "A person is wearing a **big** hat ."

Le modèle prédit bien le label neutre.

Le modèle reste cohérent dans tous les cas.

Exemple 3 :

label : inférence

P = " A person wearing a straw hat , standing outside working a steel apparatus with a pile of coconuts on the ground . "

H = " A person is near a pile of coconuts . "

Explication humaine :

Les mots importants à associer sont : "person" dans les 2 phrases, "pile of coconuts" dans les 2 phrases.

En plus de cela, le fait qu'aucun mot de l'hypothèse ne vienne contredire ou ne vienne ajouter de l'information dans la prémisse montre l'inférence.

Interprétation du modèle :

Le modèle prédit bien une inférence.

On va voir si le modèle se sert de la partie " wearing a straw hat , standing outside working a steel apparatus " de la prémisse.

- Cas 1 : P = "A person with a pile of coconuts on the ground ." ; H = " A person is near a pile of coconuts . "

Le modèle prédit bien une inférence, donc il se sert bien des bons mots de la prémisse.

On va essayer de piéger le modèle :

- Cas 2 : P = " A person wearing a straw hat , standing outside working a steel apparatus with a pile of coconuts on the ground . " (prémisse inchangée) ; H = "A person is **far away from** a pile of coconuts ."

Le modèle prédit bien une contradiction. On constate que le modèle a bien compris que les noix de coco sont proches de la personne, et n'associe pas bêtement "a pile of coconuts" dans les deux phrases pour prédire une inférence.

Exemple 4 :

label : inférence

P = " A man reads the paper in a bar with green lighting . "

H = " The man is inside . "

Explication humaine :

Les mots importants à associer sont : "man" dans les deux phrases, "in a bar" et "inside" car "bar" fait référence à un lieu fermé, donc c'est un intérieur, ce qui donne le lien avec "inside".

Interprétation du modèle :

Le modèle prédit bien une inférence et est sûr de lui.

- Cas 1 : P = "in a bar" ; H = "inside"

Le modèle prédit bien une inférence.

- Cas 2 : "A man reads the paper in a **pool** with green lighting ." ; H = " The man is inside . "

Le modèle prédit une inférence au lieu d'une contradiction (mais il hésite fortement avec la contradiction). Pourtant, en faisant P = "in a pool" et H = "inside", le modèle prédit une contradiction : il sait donc que "in a pool" n'implique pas "inside".

Peut-être que le modèle ne regarde pas " in a bar ".

- Cas 3 : P = "A man reads the paper with green lighting ." (on a retiré " in a bar ") ; H = " The man is inside . "

Le modèle prédit une inférence au lieu d'une neutralité : il ne regarde donc pas "in a bar".

- Cas 4 : P = "A man reads the paper with green lighting ." ; H = "The **dog** is inside ."

Le modèle prédit une contradiction, il sait que l'on ne parle plus de la même entité (man et dog).

- Cas 5 : P = "reads the paper" ; H = "inside"

Le modèle prédit une inférence. Peut-être qu'il fait sa prédiction en fonction du cas 5 ?

- Cas 6 : P = "A man with green lighting ." ; H = "The man is inside ."

Le modèle prédit une inférence.

- Cas 7 (de pair avec le cas 6) : P = "green lighting" ; H = "inside"

Le modèle prédit une contradiction.

donc le modèle se sert beaucoup de "man" comme nous le pensions, cependant il associe inférence à "inside" à partir de plusieurs concept : "reads the paper" et "in a bar".

Exemple 5 :

label : contradiction

P = " A couple walk hand in hand down a street . "

H = " A couple is sitting on a bench . "

Explication humaine :

Les mots importants à associer sont : "couple" dans les deux phrases pour prouver que l'on parle de la même entité, "walk" et "sitting" qui se contredisent, "street" et "bench" pour prouver que ce n'est pas le même lieu.

Interprétation du modèle :

Le modèle prédit bien une contradiction.

- Cas 1 : P = " walk " ; H = " sitting "

Le modèle prédit une contradiction.

- Cas 2 : P = " street " ; H = " bench "

Le modèle prédit une contradiction.

- Cas 3 : P = " A couple walk hand in hand down a street . " (prémisse inchangée) ; H = " A couple is sitting on a **street** . "

Le modèle prédit une contradiction. Peut-être qu'il se sert uniquement de " walk " et " sitting ".

- Cas 4 : P = " A couple walk hand in hand down a street . " (prémisse inchangée) ; H = " A couple is **walking** on a bench . "

Le modèle prédit une inférence au lieu d'une contradiction : il se sert donc de " walk " et " sitting " pour sa prédiction.

Exemple 6 :

label : neutral

P = " One tan girl with a wool hat is running and leaning over an object , while another person in a wool hat is sitting on the ground . "

H = " A man watches his daughter leap "

Explication humaine :

Les mots importants à associer sont : "girl" et "daughter" car il est possible que la fille de la prémisse soit la fille de l'homme dans l'hypothèse, "person" et "man" (ils ont un sens proche puisqu'un homme est une personne) car il est possible que la personne assise dans la prémisse soit le père de la fille, "leaning over" et "leap" car il est possible que l'on fasse un saut si l'on trébuche sur un objet.

Comme il n'y a pas d'informations sur la relation entre la fille qui tombe et la personne assise dans la prémisse, le label est neutre. Ainsi "daughter" est le mot le plus important de tous (il n'y a pas de mot tel que "dad" dans la prémisse).

Interprétation du modèle :

Le modèle prédit une contradiction au lieu d'une neutralité.

- Cas 1 : P = " one tan **boy** with a wool hat is running and leaning over an object , while another person in a wool hat is sitting on the ground . " ; H = "daughter"

Le modèle prédit une contradiction. Il a l'air de bien associer " girl " et " daughter ".

- Cas 2 : P = "one tan girl" ; H = "daughter"

Le modèle prédit une neutralité. Ce qui est correct.

Le modèle sait donc qu'il est possible que " one tan girl " soit " daughter ". On va voir s'il sait que " another person " et " man " peuvent être la même entité.

- Cas 3 : P = "another person" ; H = "man"

Le modèle prédit neutre, donc c'est bien le cas.

- Cas 4 : P = "sitting" ; H = "watches"

Le modèle prédit une inférence. Il sait donc que l'homme assis est en train de regarder.

- Cas 5 : P = "leaning" ; H = "leap"

Le modèle prédit une contradiction. Le problème vient sûrement de là.

- Cas 6 : P = "leaning over an object" ; H = "leap"

Le modèle prédit une contradiction.

- Cas 7 : P = " One tan girl with a wool hat is running and leaning over an object , while another person in a wool hat is sitting on the ground . " (prémisse inchangée) ; H = "leap"

Le modèle prédit une contradiction.

La principale action de la fille, "leap", est contradictoire avec "leaning" (cas 5 et 6) qui est l'action de la fille dans la prémisse, ainsi qu'avec toute la prémisse (cas 7).

Même si "another person"/"man" est neutre (cas 3) ainsi que "one tan girl"/"daughter" (cas 2), si toutes les actions contredisent "leap", alors le modèle choisi la contradiction.

Exemple 7 :

label : entailment

P = " A blond-haired doctor and her African american assistant looking threw new medical manuals . "

H = " A doctor is looking at a book "

Explication humaine :

Les mots importants à associer sont : "doctor" dans les deux phrases (pour prouver que c'est la même personne), "looking" dans les deux phrases, "manuals" et "book" qui veulent dire la même chose (synonymes).

Interprétation du modèle :

Le modèle prédit bien une inférence. Mais il hésite avec les autres labels.

- Cas 1 : P = "her African american assistant looking threw new medical manuals ." (" A blond-haired doctor and " retiré) ; H = " A doctor is looking at a book " (hypothèse inchangée)

Le modèle prédit une contradiction car il ne voit plus le mot doctor (ou d'autre mot se rapportant à doctor) dans la prémisse.

- Cas 2 : P = " A blond-haired doctor and her African american assistant looking threw new medical manuals . " (prémisse inchangée) ; H = "A doctor is looking at a **dog**"

Le modèle prédit une contradiction : il fait bien la différence entre " book " et " dog ", donc il arrive bien à relier " manuals " et " book ".

Exemple 8 :

label : neutral

P = " A man playing an electric guitar on stage . "

H = " A man is performing for cash . "

Explication humaine :

Les mots importants à associer sont : "man playing" et "man is performing" pour dire que c'est la même personne qui fait la même action, "for cash" dans l'hypothèse car aucune information n'est donnée là dessus dans la prémisse.

Il n'y a aucun mot ayant un sens proche avec l'argent dans la prémisse, donc c'est neutre.

Interprétation du modèle :

Le modèle prédit bien une neutralité.

- Cas 1 : P = "A man playing" (" an electric guitar on stage . " retiré) ; H = "A man is performing" (" for cash . " retiré)

Le modèle prédit bien une inférence : " playing " et " is performing " ont la même signification

- Cas 2 : P = " A man playing an electric guitar on stage . " (prémisse inchangée) ; H = " for cash . " (" A man is performing " retiré)

Le modèle prédit bien une neutralité : c'est bien "for cash" qui donne la neutralité donc le modèle applique les explications humaines.

Exemple 9 :

label : contradiction

P = " A man playing an electric guitar on stage . "

H = " A man playing banjo on the floor . "

Explication humaine :

Les mots importants à associer sont : "man playing" dans les deux phrases pour dire que c'est la même personne qui fait la même action, "electric guitar" et "banjo" qui ne sont pas les mêmes instruments, "on stage" et "on the floor" qui ne sont pas les mêmes endroits.

Interprétation du modèle :

Le modèle prédit bien une contradiction.

- Cas 1 : P = "A man playing an electric guitar" (" on stage " retiré) ; H = "A man playing banjo ." (" on the floor " retiré)

- Cas 2 : P = "electric guitar" ; H = "banjo"

Le modèle prédit bien une contradiction dans le cas 1 et le cas 2 : un banjo n'est pas une guitare électrique.

- Cas 3 : P = "A man playing on stage ." (" an electric guitar " retiré) ; H = "A man playing on the floor ." (" banjo " retiré)

Le modèle prédit une inférence au lieu d'une contradiction.

C'est donc banjo et electric guitar qui donne la contradiction.

- Cas 4 : P = "on stage" ; H = "on the floor"

Le modèle prédit ici une contradiction.

Dans le cas 3, "A man playing" prend le label inférence et c'est lui qui l'emporte (plus de mot en faveur de l'inférence que de mot en faveur à la contradiction).

Dans le cas de base, il y a beaucoup de mot en commun entre la prémisse et l'hypothèse, mais " banjo " et " electric guitar " sont considérés comme contradictoire (cas 1 et 2) ainsi que " on stage " et " on the floor " (cas 4).

Exemple 10 :

label : contradiction

P = " A land rover is being driven across a river . "

H = " A sedan is stuck in the middle of a river . "

Explication humaine :

Les mots importants à associer sont : "land rover" et "sedan" car ce sont 2 voitures (pour dire que l'on parle de la même voiture), "is being driven" et "stuck" car ils se contredisent (si la voiture est conduite ça implique qu'elle bouge, or avec "stuck" cela montre qu'elle ne bouge pas), " river " dans les deux phrases pour montrer que l'on parle du même endroit.

Interprétation du modèle :

Le modèle prédit bien une contradiction.

- Cas 1 : P = "A land rover is being driven ." (" across a river " retiré) ; H = "A sedan is stuck ." (" in the middle of a river " retiré)

Le modèle prédit bien une contradiction.

- Cas 2 : P = " A land rover is being driven across a river . " (prémisse inchangée) ; H = "A sedan is **moving** in the middle of a river . "

Le modèle prédit bien une inférence.

C'est donc le couple " is being driven " / " stuck " qui donne la contradiction.

Exemple 11 :

label : inférence

P = " A land rover is being driven across a river "

H = " A vehicle is crossing a river . "

Explication humaine :

Les mots importants à associer sont : "land rover" et "vehicle" car une land rover est un véhicule, "across a river" et "crossing a river" qui veulent dire la même chose.

Interprétation du modèle :

Le modèle prédit bien une inférence.

- Cas 1 : P = "land rover" ; H= "vehicule"

Le modèle prédit bien une inférence : une land rover est un véhicule.

- Cas 2 : P = "driven" ; H= "vehicule"

Le modèle prédit bien une inférence : " driven " est fortement lié à "vehicule".

- Cas 3 : P = " A land rover is being driven across a river " (prémisse inchangée) ; H = "A vehicle is crossing a **house**"

Le modèle prédit bien une contradiction : " river " et " house " se contredisent car " river " est un extérieur, alors que " house " est un intérieur.

- Cas 4 : P = " A land rover is being driven across a river " (prémisse inchangée) ; H = "A vehicle is crossing a **mountain**"

Le modèle prédit une inférence au lieu d'une contradiction. " river " et " mountain " sont deux extérieurs, le modèle n'y voit pas de contradiction.

- Cas 5 : P = "river" ; H = "mountain"

Le modèle prédit une inférence au lieu d'une contradiction. Cela prouve la théorie du cas 4.

Exemple 12 :

label :inférence

P = " A land rover is being driven across a river . "

H = " A Land Rover is splashing water as it crosses a river . "

Explication humaine :

Les mots importants à associer sont : "land rover" dans les deux phrases, "across a river" et "crosses a river" qui veulent dire la même chose, " river " et "splashing water".

Interprétation du modèle :

Le modèle prédit neutre au lieu d'inférence.

- Cas 1 : P = " A land rover is being driven across a river " (prémisse inchangée) ; H = "A Land Rover as it crosses a river ." ("splashing water" retiré)

Le modèle prédit bien une inférence : "splashing water" donne le label neutre, ce qui semble assez logique puisque le modèle ne peut pas savoir qu'il y a des éclaboussures.

- Cas 2 : P = " A land rover is being driven across a river " (prémisse inchangée) ; H = "A Land Rover water as it crosses a river ." (" splashing " retiré)

Le modèle prédit une neutralité : "splasing" donne le label neutre (ce mot n'a aucun lien avec les mots de la prémisse)

- Cas 3 : P = "river" ; h = "splashing water"

Le modèle prédit neutre : ça confirme ce que l'on pensait dans le cas 2.

Exemple 13 :

label : neutral

P = " An old man with a package poses in front of an advertisement . "

H = " A man poses in front of an ad for beer . "

Explication humaine :

Les mots importants à associer sont : "advertisement" et "ad for beer" respectivement dans la prémisse et l'hypothèse.

On parle ici d'une publicité, mais il n'y a aucune information supplémentaire sur la publicité en question. C'est possible que cela concerne de la bière. "for beer" rajoute de l'information.

Le fait qu'il y ait un point après "advertisement" dans la prémisse prouve qu'il n'y a pas d'infos supplémentaires. Donc le label est neutre.

Interprétation du modèle :

Le modèle prédit bien neutre.

"for beer" doit être la cause de cette prédiction donc on va l'enlever pour essayer d'avoir une inférence.

- Cas 1 : P = " An old man with a package poses in front of an advertisement . " (prémisse inchangée) ; H = "A man poses in front of an ad" (" for beer " retiré)

Le modèle prédit une inférence comme on le souhaitait : comme on le pensait, "for beer" rajoute de l'information qui n'existe pas dans la prémisse, donc le modèle prédit neutre.

Quand on l'enlève, la prémisse et l'hypothèse ont le même sens donc ça donne une inférence. Le modèle reste logique dans ce cas-là.

On va tenter de voir si, même en changeant beer, on a le label neutre. On va changer " beer " par "men" (car on parle d'un homme dans les phrases) :

- Cas 2 : P = " An old man with a package poses in front of an advertisement . " (prémisse inchangée) ; H = "A man poses in front of an ad for **men** ."

Le modèle prédit inférence au lieu de neutre : il voit "man" dans la prémisse, donc il lui associe "men" de l'hypothèse.

On va tenter de voir si ce phénomène se crée avec n'importe quel mot. On illustre cela avec le cas 3 et 4 ci-dessous :

- Cas 3 : P = "An old **beer** with a package poses in front of an advertisement . " H = " A man poses in front of an ad for beer . " (hypothèse inchangée)

- Cas 4 : P = "An old **beer** with a package poses in front of an advertisement . " H = "A man poses in front of an ad for **beers** ."

Le modèle prédit neutre dans ces deux cas. On ne retrouve pas le même phénomène, on va essayer un autre cas qui se rapproche plus du cas 2 : on va remplacer "man" par "beer" et "men" par "beers".

- Cas 5 : P = "An old **beer** with a package poses in front of an advertisement . " H = "A **beer** poses in front of an ad for **beers** ."

Le modèle prédit neutre. On ne retrouve pas le même phénomène que dans le cas 2. Cela peut s'expliquer par le fait que le modèle ait vu "man"/"men" dans beaucoup d'inférence.

On va tester un dernier phénomène : est ce que le modèle va prédire une contradiction si on remplace "for beer" dans l'hypothèse par "for women" ? Normalement il ne devrait pas, mais puisqu'il a prédit inférence dans le cas 2, alors il va sûrement prédire contradiction ici.

- Cas 6 : P = " An old man with a package poses in front of an advertisement . " (prémisse inchangée) ; H = "A man poses in front of an ad for **women** ."

Le modèle prédit une contradiction au lieu de neutre, mais ce n'est pas une surprise. Comme on le pensait, il a mis en relation "women" de l'hypothèse et "man" de la prémisse et a donné une contradiction.

- Cas 7 : P = "An old **woman** with a package poses in front of an advertisement ." ; H = "A man poses in front of an ad for **women** ."

Le modèle prédit une contradiction comme on le souhaitait. Il met donc en relation "man" dans la prémisse et dans l'hypothèse, pour savoir de qui l'on parle.

Au final, le modèle avait prédit neutre grâce à " man " dans les deux phrases car on parle de la même personne (la possibilité rentre en jeu) et grâce à " beer " car il ne contredit aucun mot de la prémisse. Le modèle n'associe pas vraiment les bons mots ensemble.