

# Simple Dialogue System with AUDITED (Supplementary Material)

Yusuf Tas  
Piotr Koniusz  
<http://users.cecs.anu.edu.au/~koniusz>

Data61/CSIRO, Australia &  
the Australian National University

## 1 Examples of Dialogues

Below, we present three randomly selected dialogues. Responses generated by our system, denoted by (*our\_pred*), appear to make fewer mistakes and be more concise.

### Example I:

User: *Show me something like the one in 4th image.*

Agent: *The similar looking ones are img\_201, img\_320, img\_513.*

User: *show me something that will pair well with the 2nd image?*

Target\_A: *It can go well with black colors sports shoes.*

MMD\_pred: *It can go well with regular style, regular fit, solid and regular style, solid and regular.*

Our\_pred: *It can go well with [casual fitted footwear](#).*

### Example II:

User: *and what about for the 1st image?*

Agent: *Image from the front, back and left directions respectively*

User: *Which of the celebrities usually wear watch similar looking to the one in the 2nd image?*

Target\_A: *Celebrities cel\_278, cel\_255 and cel\_444 endorse this kind of [watch](#).*

MMD\_pred: *Celebrities cel\_30 and cel\_252 endorse this type of [handbag](#).*

Our\_pred: *Celebrities cel\_60 and cel\_255 endorse this kind of [watch](#).*

### Example III:

Agent: *May I please ask your gender.*

User: *Female, I am a 15 yr old female.*

Agent: *Ok, this is helpful.*

Target\_A: *I'll just take a few seconds to go through my catalog of items.*

MMD\_pred: *Absolutely. Please wait a bit while I search through my catalog quickly.*

Our\_pred: *Good! Just wait a few seconds while I go through my item catalog.*

## 2 Cross-validation of NESAs

Figures 1, 2 3 evaluate the performance of NESAs model with respect to its parameters given the validation split.

In Figure 1, the bandwidth of the RBF kernel  $\sigma$  value is fixed to 0.5. We vary the number of nearest neighbors  $\eta \in \{2, 3, 4, 5, 6, 7\}$  and the number of leading eigenvectors  $\eta' \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ .

In Figure 2, number of eigenvectors is fixed to 2, that is  $\eta' = 2$ . We cross-validate the number of nearest neighbors  $\eta \in \{2, 3, 4, 5, 6, 7\}$  and the RBF bandwidth set to  $\sigma \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ .

In Figure 3, number of nearest neighbors  $\eta$  is fixed to 4. We cross-validate the number of leading eigenvectors  $\eta' \in \{0, 1, 2, 3, 4\}$  and the RBF bandwidth  $\sigma \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ .

As previously indicated, the best  $\eta' \approx \eta$ . The best  $\sigma = 0.5$  which we use across all our experiments. Moreover, as  $\eta' \approx \eta = 4$ , this indicates that our NESAs can create a rich visual context for target images (much better context than directly forcing target images to be close to their nearest neighbors in DeepFashion). Moreover, NESAs outperform NEHA (Figure 8, main submission).

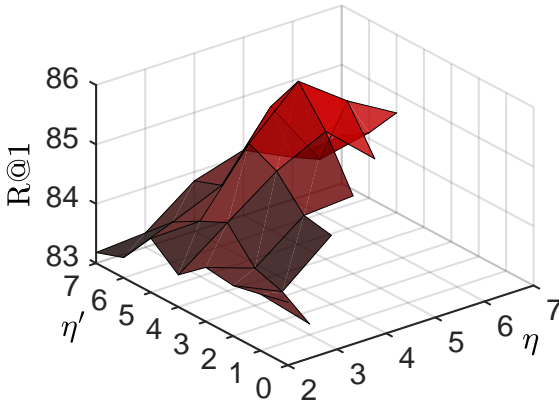


Figure 1: Performance (R@1) w.r.t. the number of nearest neighbors  $\eta$  and the number of leading eigenvectors  $\eta'$  on NESAs ( $\sigma=0.5$ ).

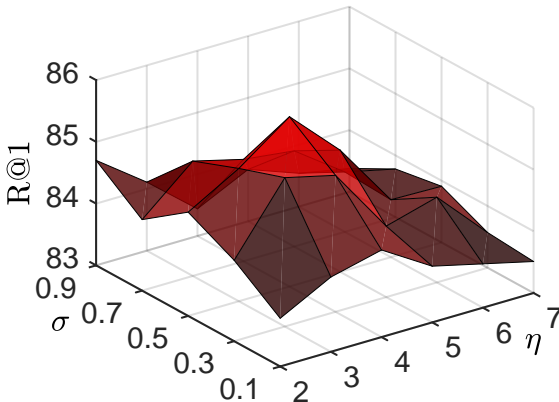


Figure 2: Performance (R@1) w.r.t. the number of nearest neighbors  $\eta$  and the RBF bandwidth  $\sigma$  on NESAs ( $\eta'=2$ ).

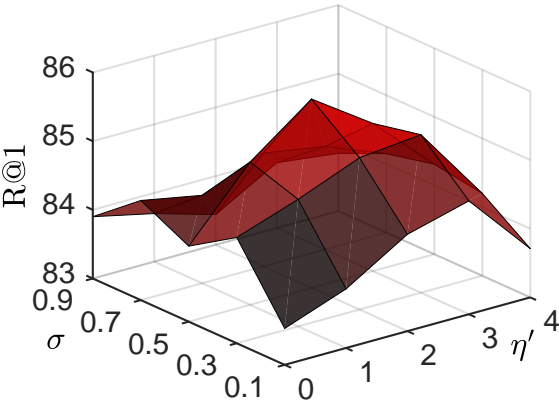


Figure 3: Performance (R@1) w.r.t. the number of leading eigenvectors  $\eta'$  and the RBF bandwidth  $\sigma$  on NESAs ( $\eta = 4$ ).